Recurrences and Legendre Transform

Volker Strehl
Institut für Mathematische Maschinen und Datenverarbeitung
(Informatik I)
Universität Erlangen-Nürnberg, Germany

Abstract

A binomial identity ((1) below), which relates the famous Apéry numbers and the sums of cubes of binomial coefficients (for which Franel has established a recurrence relation almost 100 years ago), can be seen as a particular instance of a Legendre transform between sequences. A proof of this identity can be based on the more general fact that the Apéry and Franel recurrence relations themselves are conjugate via Legendre transform. This motivates a closer look at conjugacy of sequences satisfying linear recurrence relations with polynomial coefficients. The rôle of computer-aided proof and verification in the study of binomial identities and recurrence relations is illustrated, and potential applications of conjugacy in diophantine approximation are mentioned.
This article is an expanded version of a talk given at the 29. meeting of the *Séminaire Lothringien de Combinatoire*, Thurnau, september 1992.

1 Introduction

In this article I will discuss some general aspects related to the following beautiful binomial identity:

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

An interesting aspect of this identity is the fact that it relates the famous *Apéry numbers* $a_n := \sum_k \binom{n}{k}^2 \binom{n+k}{k}^2$, which played an important rôle in R. Apéry's original proof of the irrationality of $\zeta(3)$ (see the entertaining and instructive article [20] by A. v.d. Poorten for an account of this), with the sums of cubes of binomial coefficients $f_n := \sum_k \binom{n}{k}^3$, for which J. Franel had found recurrences long ago ([6], see [4], [8], [11], [5]), and which for that reason will be called *Franel numbers* in this article.

This identity was brought to my attention early in 1992 as a conjecture by my colleague B. Voigt from Bielefeld. The origin of it lies in the following question put forward by A. L. Schmidt from Copenhagen:

Let numbers c_k $(k \geq 0)$, independent of n, be defined by

$$\sum_{k=0}^{n} \binom{n}{k}^2 \binom{n+k}{k}^2 = \sum_{k=0}^{n} \binom{n}{k} \binom{n+k}{k} c_k \quad (n \geq 0)$$

[It is easy to see (by induction) that these rational numbers are uniquely determined.]

Is it then true that these numbers are always integers?

[This is not obvious from the definition, but an evaluation of the first few hundred values clearly suggests that.]

It was soon observed by W. Deuber, W. Thumser and B. Voigt that the beginning of the sequence

$$(c_k)_{k\geq 0} = (\,1\,,\,2\,,\,10\,,\,56\,,\,346\,,\,2252\,,\,15184\,,\,104960\,,\,739162,\ldots)$$

is the same as the beginning of the sequence of the Franel numbers $f_k = \sum_{j=0}^{k} \binom{k}{j}^3$. From this observation the above conjecture (1) was formulated. Since I could not spot the conjectured identity in the relevant literature (e.g. [12], [8]), I tried to find a proof for myself. In fact, there are now several different proofs available, none of them really "trivial", and I will briefly mention below these diverse approaches that I worked out. The ordering corresponds to the chronological ordering in which these proofs were discovered. A detailed discussion of this affair will be given in [18].

1. Identity (1) is hidden in a classical result, due to W.N. Bailey, from the theory of hypergeometric functions. In fact, it can be pulled out of Bailey's bilinear generating function for the Jacobi polynomials ([2], [17], [19]) via a kind of diagonalization. The special case where the Jacobi polynomials degenerate into Legendre polynomials is sufficient for our identity, but the same trick can be played in the general situation, so that (1) turns out to be a degenerate case belonging to a family of binomial identities with two free parameters.

2. As has been shown by myself in [19], Bailey's bilinear generating function can be proved from a suitable combinatorial interpretation. This is somewhat involved, and on the way of proving the full bilinear generating function I established a certain family of binomial identities. It turned out that these already include (1) (up to routine transformations) and its generalization mentioned before. This answers positively the question about a possible combinatorial proof of (1), but such a proof is more complicated than the rather simple appearance of this identity might suggest.

3. By using the method of Legendre inverse pairs, as exposed in Riordan's book [12], the problem can be transformed into an equivalent one to which D. Zeilberger's method of constructing recurrence operators for definite hypergeometric sums (in the case of single sums) can be applied. This gives a "computer-supported" proof, which can be turned into a "conventional" one - once the operators are known.

4. The desired result can be obtained by a series of applications of hypergeometric transforms (D. Stanton, Ch. Krattenthaler and G. Andrews have provided comments on that approach). Once found, such a proof is routine to check - finding the right track is the problem!

5. The Wilf-Zeilberger-method (see remarks below) for multiple hypergeometric sums can be applied directly to the identity. Doron Zeilberger provided such a completely "machine-made" proof in june 1992.

6. Using the known recurrences for the Apéry numbers $a_n = \sum_k \binom{n}{k}^2 \binom{n+k}{k}^2$ and the Franel numbers $f_n = \sum_k \binom{n}{k}^3$ one can give a short (partially computer supported) proof of (1), using no sophisticated algorithms, but only standard simplification capabilities of any reasonable computer algebra program. Such a proof is outlined in the next section, and indeed, this kind of proof was the motivation for writing the present note. Let me remark that a proof similar to this one was found independently by A. L. Schmidt.

The following sections are organized as follows.

After illustrating the use (neither the principles, nor the theory behind) of available implementations of algorithms for a computer-aided treatment of binomial sums and sequences satisfying linear recurrence relations with polynomial coefficients, I will present a short proof of (1) which is based on the (known) recurrence relations of the Apéry and Franel numbers. The crucial fact is this: identity (1) claims that Franel's and Apéry's sequences are related via *Legendre transform.* This concept has been used, with number-theoretic applications in mind, by A. L. Schmidt, see [14], [15], [16], and, in particular, his most recent article is closely related to the present one, and some of his results will be mentioned below. This turns out to be an immediate consequence of the stronger fact that the linear recurrences (of second order, with polynomial coefficients) generating those sequences are themselves *conjugates* via Legendre transform in a sense to be made precise below. Once guessed correctly, a proof of this conjugacy relation is remarkably simple (relative to a routine use of a machine-based simplification procedure).

In section 3 some notation will be developed in order to deal with this approach from a more general perspective.

The situation of (reduced) Legendre transform is then studied in some detail in section 4, and the question under which condition (and if so: how) recurrence relations are transformed into recurrence relations via conjugation is answered completely in a particular case. It should become clear, however, that similar results could be obtained by the same approach in related and more general situations. Thus these results should be taken as exemplary - no complete treatment was intended here, just an outline of how one may proceed.

In the final section I present some consequences in the field of diophantine approximation which are directly related to identity (1) and the conjugacy relation behind it. These results are taken from A. L. Schmidt's recent article [16] mentioned above - here they should be taken as a motivation showing that the study of binomial (or hypergeometric) identities and recurrences related to them might lead to interesting applications in number theory.

2 A short proof of identity (1)

Two basic facts are used for a proof of (1) as presented here:

- [Franel, 1895] The Franel numbers $f_n = \sum_k \binom{n}{k}^3$ satisfy the following second-order recurrence with polynomial coefficients [1]

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

- [Apéry, 1978] The Apéry numbers $f_n = \sum_k \binom{n}{k}^2 \binom{n+k}{k}^2$ satisfy the following second-order recurrence with polynomial coefficients[2]

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

Note that even if one *knows* the recurrences beforehand, it can be rather tedious to *verify* them. To illustrate this situation, let me cite (with a slight change in notation)[3] from A. v.d. Poorten's article [20]:

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

Then v.d. Poorten mentions that D. Zagier was able to provide a solution to this problem *with irritating speed* by virtue of a technique called *creative telescoping*. Actually, this technique, combined with the powerful algorithm for indefinite hypergeometric summation due to W. Gosper [7], is at the heart of D. Zeilberger's method. Today, finding (and proving!) recurrences like (2) and (3) is a routine application of Zeilberger's algorithm for simple terminating hypergeometric sums, in fact a matter of very few seconds (in these cases! In a case like the one mentioned above in 5., i.e. directly verifying that the right hand side of (1) satisfies the same recurrence as the Apéry numbers do, application of Zeilberger's algorithm for *hypergeometric multisums* leads to nonnegligible computing time, even on fast machines[4]. To illustrate this, I include input and output from an implementation

[1] Franel has also found a similar recurrence for $\sum_{k=0}^n \binom{n}{k}^4$. It took quite a while until corresponding recurrence relations for higher exponents were found, see the recent articles by Perlstadt [11] and Cusick [5].

[2] Apéry has also found a similar recurrence for $\sum_{k=0}^n \binom{n}{k}^2 \binom{n+k}{k}$. See the article by van der Poorten [20] and an article [1] by R. Askey and J. Wilson for illuminating remarks from the hypergeometric viewpoint and a generalization.

[3] Actually, v.d. Poorten writes about two solutions of the recurrence, the other one, call it $\{ b_n \}$ specified by initial values $b_0 = 0$ and $b_1 = 6$. This second solution will only appear in the last section of this article. Note that the rôle of a and b has been interchanged w.r.t. [20].

[4] D. Zeilberger's first direct verification of (1) took about 1500 sec.

of Zeilberger's algorithm which was programmed by J. Hornegger [9] in the context of the AXIOM[5] computer algebra system.

Recurrences for $\sum_{k=0}^{n} \binom{n}{k}^e \quad (e = 1, 2, 3)$

```
k1 : Symbol := k
p1 : SMP(FRAC INT, Symbol)      := n
p2 : SMP(FRAC INT, Symbol)      := k
arg: FRAC SMP(FRAC INT, Symbol) := 1
```

Exponent $e = 1$:

```
(69)  r1:= findrec(bin(p1, p2, n1, k1, 1), 1, arg, 1, n1, k1)

  (69)  - B + 2, 1

Type: Union(List FRAC SMP(FRAC INT,Symbol),...)
             Time: 0.08 (IN) + 0.13 (EV) + 0.04 (OT) = 0.25 sec
```

The result consists of the recurrence operator (written as a polynomial in the shift operator B and n) and the *certificate* provided by the method (plus informations about typing and timing).

Exponent $e = 2$:

```
(71)  r2:= findrec(bin(p1, p2, n1, k1, 2), 1, arg, 1, n1, k1)

  (71)  (B - 4)n + B - 2, - 3n + 2k - 1

Type:  ...
            Time: 0.08 (IN) + 0.31 (EV) + 0.05 (OT) = 0.44 sec
```

This result expresses the fact that $\sum_{k=0}^{n} \binom{n}{k}^2 = \binom{2n}{n}$, because both sides are annihilated by the recurrence operator.

Exponent $e = 3$:

```
(72)  r3:= findrec(bin(p1, p2, n1, k1, 3), 1, arg, 1, n1, k1)

  (72)

   2            2       2                    2
 (B  - 7 B - 8)n  + (4 B  - 21 B - 16)n + 4 B  - 16 B - 8,

              5            4         2               3
        - 14n  + (27k - 75)n  + (- 18k  + 111k - 161)n
       +
           3      2                  2      3      2
        (4k  - 54k  + 171k - 173)n  + (8k  - 54k  + 117k - 93)n
       +
          3      2
        4k  - 18k  + 30k - 20
```

[5] *AXIOM* is a trademark of NAG, Numerical Algoriths Group Ltd.,Oxford, England.

```
    /
        3              2      2               3     2
       n  + (- 3k + 3)n  + (3k  - 6k + 3)n - k  + 3k  - 3k + 1

Type:  ...

         Time: 0.08 (IN) + 13.30 (EV) + 0.38 (OT) +4.15 (GC) = 17.91 sec

(73)  factorCoefficients r3.1

                2 2      2        16            2
   (73)  (n + 2) B  - 7(n  + 3n + --)B - 8(n + 1)
                                  7

Type: SUP Factored SMP(FRAC INT,Symbol)
         Time: 0.01 (IN) + 0.11 (EV) + 0.09 (OT) = 0.21 sec
```

This is Franel's recurence. The certificate in factorized form:

```
(74)  factorCertifyingFunction r3.2

   (74)
                 2
      - 14(n + 1)
   *
     3      27     47  2    9  2   57     53     2  3   9  2  15     10
    (n  + (- -- k + --)n  + (- k  - -- k + --)n - - k  + - k - -- k + --)
             14     14       7      14     14     7      7      7      7
 /
               3
    (n - k + 1)
Type: FRAC Factored SMP(FRAC INT,Symbol)
        Time: 0.04 (IN) + 0.78 (EV) + 0.48 (OT) + 1.50 (GC) = 2.80 sec
```

Now for Apéry's recurrence (omitting the certificate):

```
a1: SMP(FRAC INT, Symbol):= n
a2: SMP(FRAC INT, Symbol):= k
a3: SMP(FRAC INT, Symbol):= a1 + a2

r:= recurrence  findrec(bin(a1, a2, n1, k1, 2),
     bin(a3, a2, n1, k1, 2), 1, 2, n1, k1)

  3     2            2         3       2
(n  + 6n  + 12n + 8)B  + (- 34n  - 153n  - 231n - 117)B +
       3     2
      n  + 3n  + 3n + 1

Type: SUP SMP(FRAC INT, Symbol)
        Time: 0.20 (IN) + 13.71 (EV) + 0.23 (OT) = 14.14 sec
```

Here I will not give any further explanations w.r.t. Zeilberger's method. The interested reader may look into the numerous articles available (e.g. [25], [26],

[21], [22], [24], [23], [3]), [10], [9]) describing the method, the algorithms, and their applications.

As a side remark: I would like to take the occasion to point out another approach, experimental in character, which can be helpful in situations similar to the one discussed here. There is now a nice set of tools available with which given the first few coefficients of a generatin function, one may *cleverly guess* the generating function (if it has a "nice" closed form), or a linear differential equation satisfied by it, or a linear recurrence relation satisfied by its coefficients. The mathematical model behind it is that of *holonomic* functions of one variable (or equivalently: D-finite power series, or P-recursive sequences), with important special cases such as rational and algebraic generating functions. Even if these tools only provide *guesses* and not *proofs* in the strict sense, they often can be used to obtain enough information about a problem so that a strict solution becomes feasible. These tools, written by B. Salvy and P. Zimmerman (INRIA, Paris) are available under the Maple[6] computer algebra system as a package named *gfun* in the Maple share library [7]. A rather comprehensive description [13] will soon be available from the authors.

As an illustration, let us *guess* the recurrences (2) and (3) using *gfun*:

```
    |\^/|      MAPLE V
._|\|   |/|_.  Copyright (c) 1981-1990 by the University of Waterloo.
 \  MAPLE  /   All rights reserved.  MAPLE is a registered trademark of
 <____ ____>   Waterloo Maple Software.
      |        Type ? for help.

> readlib(gfun):with(gfun):
```

The Apéry numbers are defined:

$$a := \left\langle n \mapsto \sum_{k=0}^{n} \binom{n}{k}^2 \binom{n+k}{k}^2 \right\rangle$$

A list of the first few values:

```
> [a(i)$i=0..10];

       [1, 5, 73, 1445, 33001, 819005, 21460825, 584307365,

       16367912425, 468690849005, 13657436403073]
```

gfun does not find a "nice" generating function:

```
> guessgf([a(i)$i=0..15],t);
                                FAIL
```

Eleven initial values are not sufficient for the recurrence:

[6] *Maple* is a trademark of Waterloo Maple Software, Waterloo, Canada.

[7] The Maple share library be accessed using anonymous ftp from `daisy.uwaterloo.ca` in Waterloo (Canada) or `neptune.inf.ethz.ch` in Zurich (Switzerland).

```
> listtorec([a(i)$i=0..10],A(n));
```

FAIL

but sixteen initial values are:

```
> listtorec([a(i)$i=0..15],A(n));
```

$$[\ \{n^3 A(n) + \left(-5 - 27\,n - 51\,n^2 - 34\,n^3\right) A(n+1) + \left(1 + 3\,n + 3\,n^2 + n^3\right) A(n+2), A(1) = 1, A(2) = 5\}\ ,\ ogf\ ]$$

We can get a differential equation from the recurrence:

```
> rectodiffeq("[1],A(n),aa(z));
```

$$\begin{aligned}[\ & \mathrm{D}(aa)(0) = 1, D^{(2)}(aa)(0) = 10, \\ & \left(3\,z^4 - 51\,z^3\right) \frac{d^2}{dz^2} aa(z) + \left(z + z^3 - 10\,z^2\right) \frac{d}{dz} aa(z) + (5\,z - 1)\,aa(z) \\ & -5\,A(0) z + A(0) + z^5 \frac{d^3}{dz^3} aa(z) + \left(\frac{d^3}{dz^3} aa(z)\right) z^3 - 34\left(\frac{d^3}{dz^3} aa(z)\right) z^4\ ,\ ogf\ ]\end{aligned}$$

Now we consider sums of cubes of binomial coefficients:

$$f := \left\langle n \mapsto \sum_{k=0}^{n} \binom{n}{k}^3 \right\rangle$$

```
> [f(i)$i=0..10];

[1, 2, 10, 56, 346, 2252, 15184, 104960, 739162, 5280932, 38165260]
```

Too few initial values will not lead to a differential equation:

```
> listtodiffeq([f(i)$i=0..10],F(z));
```

FAIL

```
> listtodiffeq([f(i)$i=0..15],F(z));
```

$$\begin{aligned}\{\ & (-2 - 8\,z)\,F(z) + \left(1 - 14\,z - 24\,z^2\right) \frac{d}{dz} F(z) + \left(z - 7\,z^2 - 8\,z^3\right) \frac{d^2}{dz^2} F(z), \\ & \mathrm{D}(F)(0) = 2, F(0) = 1\ \}\end{aligned}$$

From the differential equation we get the Franel recurrence:

```
> diffeqtorec(",F(z),ff(n));
```

$$\begin{aligned}\{\ & \left(-7n - 2 - 7n^2\right) \mathit{ff}(n) + \left(n^2 + 2n + 1\right) \mathit{ff}(n+1) - 8\mathit{ff}(n-1)n^2, \\ & \mathit{ff}(0) = 1, \mathit{ff}(1) = 2\ \}\end{aligned}$$

Let us now return to the proposed proof of (1). It will be presented in matrix form. To begin with, let us define a doubly infinite tridiagonal matrix $\mathbf{F} = (f_{i,j})_{i,j\geq 0}$ representing the difference operator of the Franel recurrence (2) :

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

so that

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

If we denote by $\mathbf{f} = (f_n)_{n\geq 0}$ the infinite column vector with the Franel numbers $f_n = \sum_k \binom{n}{k}^3$ as entries, then Franel's result (2) is equivalent to

$$\mathbf{F} \cdot \mathbf{f} = \mathbf{0}$$

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

Similarly, the difference operator occuring in the Apéry recurrence (3) has matrix form

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

where

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

Then Apéry's recurrence can be written as

$$\mathbf{A} \cdot \mathbf{a} = \mathbf{0}$$

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

Finally we introduce the doubly infinite, lower triangular matrix $\mathbf{P} = (p_{i,j})_{i,j\geq 0}$ of the *Legendre transform*:

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

Note that the n-th row of this matrix contains the coefficients of the Legendre polynomials P_n, if written as

$$P_n(z) = \sum_{k=0}^{n} \binom{n}{k} \binom{n+k}{k} \left(\frac{z-1}{2} \right)^k$$

Using this notation, it becomes clear that A.L. Schmidt's question mentioned in the beginning asks for the inverse image of the sequence of Apéry's numbers under the Legendre transform, and that the conjectured identity (1) claims that Apéry's sequence $\mathbf{a}$ is the image of Franel's sequence $\mathbf{f}$ under the Legendre transform, i.e. $\mathbf{a} = \mathbf{P} \cdot \mathbf{f}$. Put into matrix terms, what we would like to show is that

$$\mathbf{A} \cdot \mathbf{P} \cdot \mathbf{f} = \mathbf{0}$$

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

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

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

$$\mathbf{X} = \mathbf{D} \cdot \mathbf{P} \quad \text{and hence} \quad \mathbf{A} = \mathbf{D} \cdot \mathbf{P} \cdot \mathbf{F} \cdot \mathbf{P}^{-1} \tag{4}$$

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

It remains to *prove* this claim which not only says that the Franel and the Apéry sequences are related via the Legendre transform, but also that the associated difference operators are related in the sense of *conjugation* via Legendre transform (up to multiplication with a diagonal matrix). Interestingly, even though we have to deal with infinite matrices containing binomial coefficients, i.e. nonrational terms, the proof can be established by rational arithmetic alone. For this to see, let us write the assertion (4) to be proved in the form:

$$\mathbf{A} \cdot \mathbf{P} = \mathbf{D} \cdot \mathbf{P} \cdot \mathbf{F}$$

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

$$\begin{aligned} lhs &= a_1(i)\, p_{i-1,j} + a_0(i)\, p_{i,j} + a_{-1}(i)\, p_{i+1,j} \\ rhs &= (4i+2)\, [p_{i,j-1}\, f_{-1}(j-1) + p_{i,j}\, f_0(j) + p_{i,j+1}\, f_1(j+1)] \end{aligned}$$

where now i and j are treated as *variables*. Now ask for simplification of $lhs - rhs$. The Maple command

```
expand(lhs - rhs)
```

gives back an expression of considerable size. Of course, simplification could be done by hand, but this tedious task is better accomplished by your computer algebra system[8]:

[8]The subtle point here is this: the simplification algorithm must recognize the fact that even though there are binomial coefficients with symbolic entries around, the problem is essentially one of rational arithmetic because of the *hypergeometric* nature of binomial coefficients: walking along rows and columns in the Pascal triangle can locally be performed via multiplication with suitable rational functions of the locations, hence rational normalization is possible as soon as only bounded neighbourhoods have to be taken into account. Here the tridiagonal structure of the matrices $\mathbf{A}$ and $\mathbf{F}$ is important.

```
simplify(expand(lhs - rhs))
```

```
0
```

To conclude this section, let us summarize the above discussion

Theorem 1 *The Franel recurrence operator* $\mathbf{F}$ *and the Apéry recurrence operator* $\mathbf{A}$ *are Legendre conjugates in the following sense:*

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

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

The binomial identity (1) we started with is just one of the consequences of this general fact - perhaps the most interesting one. In section 5 it will be shown how this conjugacy relation can be further exploited.

Note that in the proof of the theorem we essentially used the fact that the matrices $\mathbf{A}$ and $\mathbf{F}$ are tridiagonal matrices with "polynomial values along the diagonals", and that the matrix of the Legendre transform $\mathbf{P}$ has a similar structure[9]. It is therefore natural to ask under which conditions a conjugation operation (with a nonsingular, lower triangular matrix such as $\mathbf{P}$) will transform matrices representing linear recurrence operators with polynomial coefficients into matrices of the same kind. I will not try to give a definitive answer to this question in generality. Instead, I will look into the case of the (reduced) Legendre transform and show how, at least for recurrence operators of low order and with low degree polynomial coefficients, such a question can be attacked.

3 Notation and some generalities

Let $\Phi = (\phi_i(z))_{i\in\mathbf{Z}}$ be a family of functions defined on the integers. With this family we will associate a doubly-infinite matrix $D_\Phi = (d_{i,j})_{i,j\geq 0} = (\phi_{i-j}(i))_{i,j\geq 0}$, i.e.

$$D_\Phi = \begin{pmatrix} \phi_0(0) & \phi_{-1}(0) & \phi_{-2}(0) & \dots \\ \phi_1(1) & \phi_0(1) & \phi_{-1}(1) & \dots \\ \phi_2(2) & \phi_1(2) & \phi_0(2) & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

In the previous section, we have seen the matrices $\mathbf{A}$ and $\mathbf{F}$ belonging to the Apéry and the Franel recurrence, respectively, when written in this form, are defined by *polynomial* families $(a_i(z))_{i\in\mathbf{Z}}$ and $(f_i(z))_{i\in\mathbf{Z}}$ (with three nonzero members in each case). The matrix $\mathbf{P}$ of the Legendre transform does not belong to a polynomial

[9]No quite, actually: the values "along the diagonals" are no longer polynomial, but still hypergeometric. That is why in section 4 we will pass to a related transform, called *reduced* Legendre transform, which is polynomial and thus behaves a bit nicer.

family, simply because $i \mapsto \binom{2i}{i}$ does not grow polynomially. (One might use the Γ-function in order to represent the matrix by a family of analytic functions. Instead, we will pass to a related polynomial family below).

In particular, if $\Phi = (t_i(z))_{i\in\mathbf{Z}}$ is such that $t_i = 0$ for $i < 0$ and $t_0(j) \neq 0$ for all $j \in \mathbf{N}$, i.e. D_Φ is an invertible lower triangular triangular matrix, then its inverse matrix $(D_\Phi)^{-1}$ can be written as D_Ψ where the family $\Psi = (t'_i(z))_{i\in\mathbf{Z}}$ is given by $t'_i(z) = 0$ for $i < 0$, $t'_0(j) = 1/t_0(j)$ for all $j \in \mathbf{N}$, and for $i > 0$, $(-1)^i t'_i(z)$ is given by:

$$\frac{\det\begin{pmatrix} t_1(z-i+1) & t_0(z-i+1) & 0 & 0 & \dots & 0 \\ t_2(z-i+2) & t_1(z-i+2) & t_0(z-i+2) & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ t_{i-1}(z-1) & t_{i-2}(z-1) & \dots & \dots & \dots & t_0(z-1) \\ t_i(z) & t_{i-1}(z) & t_{i-2}(z) & t_{i-3}(z) & \dots & t_1(z) \end{pmatrix}}{t_0(z-i)\cdot\ldots\cdot t_0(z)}$$

This shows in particular: if Φ is a *polynomial* family in the sense that all $t_i(z)$ are polynomials, and if in addition $t_0(z)$ is a non-vanishing constant, then the inverse family Ψ is of the same kind. Moreover, if $\deg t_i(z) \leq i$ for all $i \geq 0$, then expansion of the determinant shows that $t'_i(z)$ is a sum of terms of degree $\leq i$, hence $\deg t'_i(z) \leq i$ for $i \geq 0$.

Now take a function $g(z)$ and $\delta \in \mathbf{Z}$. The singleton family $\Gamma_{g,\delta} = (g_i(z))_{i\in\mathbf{Z}}$ is given by $g_\delta(z) = g(z)$ and $g_i(z) = 0$ for $i \neq \delta$. Then consider the family $\Lambda = (h_d(z))_{d\in\mathbf{Z}}$ given by conjugation

$$D_\Phi \cdot D_\Gamma \cdot (D_\Phi)^{-1} = D_\Lambda$$

with a 'triangular' family $\Phi = (t_i(z))_{i\in\mathbf{Z}}$ as before. It is easy to see that $h_d(z) = 0$ for $d < \delta$, and that for $d \geq \delta$ one gets

$$h_d(z) = \sum_{0\leq k\leq d-\delta} t_k(z)\cdot g(z-k)\cdot t'_{d-\delta-k}(z-k-\delta) \tag{5}$$

which can be written in determinantal form (and thus avoiding explicit reference to the inverse family $t'_i(z)$) as

$$h_d(z) = \frac{\det M(\Phi, g, \delta, d)}{t_0(z-d)\cdot\ldots\cdot t_0(z-\delta)}$$

where the square matrix $M(\Phi, g, \delta, d)$ of size $d - \delta + 1$ is given by

$$\begin{pmatrix} t_{d-\delta}(z)\cdot g(z-d+\delta) & t_{d-\delta-1}(z)\cdot g(z-d+\delta+1) & \dots & t_0(z)\cdot g(z) \\ t_1(z-d+1) & t_0(z-d+1) & \dots & 0 \\ t_2(z-d+2) & t_1(z-d+2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ t_{d-\delta}(z-\delta) & t_{d-\delta-1}(z-\delta) & \dots & t_0(z-\delta) \end{pmatrix}$$

Either of these presentations shows that for a polynomial $g(z)$ and $\delta \in \mathbf{Z}$ the family $\Lambda = (h_i(z))_{i\in\mathbf{Z}}$ obtained via conjugation with a polynomial family $\Phi = (t_i(z))_{i\in\mathbf{Z}}$ is again polynomial (and vanishes for indices $d < \delta$).

By superposition and linearity, if $\Gamma = (g_i(z))_{i\in\mathbf{Z}}$ is a family of polynomials such that $g_i(z) = 0$ for all i less than some integer K, and if Φ is as before, than conjugation

$$D_\Phi \cdot D_\Gamma \cdot (D_\Phi)^{-1}$$

is well-defined and gives a family of polynomials $\Lambda = (h_d(z))_{d\in\mathbf{Z}}$ which vanishes for $d < K$, and for $d \geq K$ its members are given by

$$h_d(z) = \sum_{K\leq\delta\leq d}\ \sum_{0\leq k\leq d-\delta} t_k(z) \cdot g_\delta(z-k) \cdot t'_{d-\delta-k}(z-k-\delta)$$

In general, if the family $\Gamma = (g_i(z))_{i\in\mathbf{Z}}$ has a finite number of nonzero terms - and hence corresponds to a linear difference operator with polynomial coefficients - the same will not be true for the members of the conjugate family $\Lambda = (h_i(z))_{i\in\mathbf{Z}}$, as the case of a singleton family $\Gamma_{g,\delta}$ clearly shows. We also expect the degree of the conjugate polynomials $h_i(z)$ to be higher than the degrees of the $g_j(z)$. Thus it will come rather as an exception if the conjugate under Φ-transform of a finite polynomial family Γ (in the above sense) is again of the same type.

In the next section we will study this situation with respect to a variant of the Legendre transform.

4 (Reduced) Legendre transform preserves degrees

We consider now the matrix $\mathbf{P} = (p_{i,j})_{i,j\geq 0}$ belonging to the Legendre transform, i.e.

$$p_{i,j} = \binom{i}{j}\binom{i+j}{j} = \binom{2j}{j}\binom{i+j}{i-j}$$

which corresponds to a family $\Pi = (p_k(z))_{k\in\mathbf{Z}}$ whose members do not behave polynomially. Since it is slightly more convenient to work with polynomials only, we will write

$$\mathbf{P} = \mathbf{Q} \cdot \boldsymbol{\Delta}$$

where

$$\mathbf{Q} = (q_{i,j})_{i,j\geq 0} \quad \text{with} \quad q_{i,j} = \binom{i+j}{i-j} \quad \text{and} \quad \boldsymbol{\Delta} = \operatorname{diag}\left(\binom{2i}{i}\right)_{i\geq 0} \tag{6}$$

The matrix $\mathbf{Q}$ belongs to the polynomial family $\Phi = (q_k(z))_{k\in\mathbf{Z}}$ where $q_k(z) = \binom{2z-k}{k}$ $(k \geq 0)$. Note that the family $\Phi' = (q'_k(z))_{k\in\mathbf{Z}}$ belonging to the inverse matrix $\mathbf{Q}^{-1}$ is given by

$$q'_k(z) = (-1)^k \left[\binom{2z}{k} - \binom{2z}{k-1}\right] = (-1)^k \binom{2z}{k}\frac{2z-2k+1}{2z-k+1} \tag{7}$$

so that $q_0(z) = q_0'(z) = 1$ and $\deg q_k(z) = \deg q_k'(z) = k$ $(k \geq 0)$. The transformation provided by the matrix $\mathbf{Q}$ and its inverse will be called *reduced Legendre transform.* Since it differs from the Legendre transform just by multiplication with the diagonal matrix Δ, it will be easy to carry over results from one situation to the other.

Let us now take singleton families $\Gamma_{g,\delta}$ as above, with polynomial $g(z)$ and $\delta \in \mathbf{Z}$, and study their reduced Legendre transform. Let $\Lambda = (h_d(z))_{d\in\mathbf{Z}}$ be the conjugate family obtained from (reduced) Legendre transform, i.e.

$$D_\Lambda = D_\Phi \cdot D_{\Gamma_{g,\delta}} \cdot (D_\Phi)^{-1}$$

Below I will sketch a proof of

Proposition 2 *With the notions introduced, one has* $\deg h_d(z) \leq \deg g(z)$ *for all* $d \geq \delta$*, and hence the same result holds by linearity for any (summable) combination of such singleton families*

The proof will only given as a sketch, technical details will be left out. Note that the assertion is robust under linear transformations of the z-variable, so free use can be made of this fact.

From (5),(6), and (7) we know that we have to evaluate

$$h_d(z) = \sum_{k=0}^{n} g(z-k)\binom{2z-k}{k}(-1)^{n-k}\binom{2(z-k-\delta)}{n-k}\frac{2(z-\delta-n)+1}{2(z-\delta)-n-k+1}$$

where $n = d-\delta \geq 0$. By expanding and linear transformation it turns out that it is sufficient to show that for $t \geq 0$ a sum like

$$\sum_{k=0}^{n}(-1)^{n-k}\,k^t\binom{z+d-k}{k}\binom{z-2k}{n-k}\frac{z-2n+1}{z-n-k+1}$$

is a polynomial of degree $\leq \min(n,t)$ in z. Now, this follows from

Lemma 3 *For each integer* $t \geq 0$

$$r_{n,t}(z,d) = \sum_{k=0}^{n}(-1)^{n-k}k(k-1)\dots(k-t+1)\binom{z+d-k}{k}\binom{z-2k}{n-k}\frac{z-2n+1}{z-n-k+1}$$

is a polynomial in z *and* d *with:*

1. $r_{n,t}(z,d) = 0$ *for* $0 \leq n < t$ *(trivially);*
2. *for* $n \geq t$, $r_{n,t}(z,d)$ *is a polynomial of total degree* n*, and its terms of highest degree are given by*

$$\frac{1}{(n-t)!}(z+d)^t \cdot d^{n-t}$$

3. the degree in z of $r_{n,t}(z,d)$ is always $\le t$, and in particular

$$r_{n,0}(z,d) = \binom{n+d-1}{n}$$

For our purposes, we are only interested in the assertion about the z-degree. The proof of this lemma can be reduced to the proof of a similar assertion for sums like

$$\sum_{k=0}^{n} (-1)^k \, k^t \binom{z+d-k}{k} \binom{z-2k}{n-k}$$

or

$$\sum_{k=0}^{n} (-1)^k \binom{k}{t} \binom{z+d-k}{k} \binom{z-2k}{n-k}$$

for nonnegative integer values of t.

Let us now consider a bit more generally sums like

$$\sum_{k=0}^{n} (-1)^k \binom{k}{t} \binom{x+\lambda\, k}{k} \binom{y+\mu\,(n-k)}{n-k} \tag{8}$$

Note that the above sums show up when specializing:

$$x = z+d \ , \ y = z-2n \ , \ \lambda = -1 \ , \ \mu = 2$$

For $t > n$ it is clear, that the sum vanishes, since every summand equals zero. For $0 < t < n$, the first t terms of (8) vanish, and this together with rewriting of binomial coefficients and linear transformation of the variables makes it possible to replace the sum by a similar one in which the range $0 \ldots n$ has been replaced by $0 \ldots n-t$. (Some care has to be taken to see that the inductive argument goes through in this situation).
The critical case which remains to be treated and which can be eventually be made responsible for the cancellation phenomenon is then the case $t = 0$. Here we can use:

Lemma 4

$$\sum_{k=0}^{n} (-1)^k \binom{x+\lambda k}{k} \binom{y+\mu(n-k)}{n-k}$$

can be written as a polynomial in $(x-y)$, provided $\lambda + \mu = 1$.

In particular, if we replace x by $z+a$ and similarly y by $z+b$, for some a, b independent of z and k, the the sum is a constant w.r.t. z.

Let us now take a tridiagonal matrix $\mathbf{G}$ associated with polynomials of second degree, i.e. $\mathbf{G} = D_\Gamma$ where $\Gamma = (g_i(z))_{i\in\mathbf{Z}}$ s. th.

$$\begin{aligned} g_{-1}(z) &= a_0 + a_1\, z + a_2\, z^2 \\ g_0(z) &= b_0 + b_1\, z + b_2\, z^2 \\ g_1(z) &= c_0 + c_1\, z + c_2\, z^2 \\ g_k(z) &= 0 \ \text{ if } \ |k| > 1 \end{aligned} \tag{9}$$

Let $\Lambda = (h_d(z))_{d \in \mathbf{Z}}$ be the conjugate family obtained from Γ via reduced Legendre transform. We know from the above considerations, that for $d \geq -1$ the polynomials $h_d(z)$ are polynomials of degree 2 if we take the coefficients $a_0, \ldots, c_2$ as indeterminates. What we are interested in is the question: *under which conditions will the matrix D_Λ again be a tridiagonal matrix, (i.e. $h_d(z) = 0$ for $|d| > 1$) ?* It should be clear that the occurence of this phenomenon is expressible as a linear condition in the indeterminates. So let us use the above formulas and calculate the first few values of $h_d(z) = \lambda_{d,0} + \lambda_{d,1}\, z + \lambda_{d,2}\, z^2$

d	$\lambda_{d,0}$	$\lambda_{d,1}$	$\lambda_{d,2}$
-1	a_0	a_1	a_2
0	$-2a_0 + a_1 - a_2$	$-4a_1 + 4a_2 + b_1$	$-6a_2 + b_2$
1	$a_0 - 5a_1 + 11a_2$	$7a_1 - 26a_2$	$17a_2 - 4b_2 + c_2$
	$+b_1 - b_2 + c_0$	$-2b_1 + 4b_2 + c_1$	
2	$12a_1 - 54a_2$	$-8a_1 + 48a_2$	$-32a_2 + 8b_2 - 2c_2$
	$-3b_1 + 9b_2$	$+2b_1 - 18b_2$	
	$2c_0 + c_1 - c_2$	$+4c_2$	
3	$-20a_1 + 170a_2$	$8a_1 - 188a_2$	$48a_2 - 12b_2 + 3c_2$
	$+5b_1 - 35b_2$	$-2b_1 + 44b_2$	
	$+3c_0 - c_1 + 7c_2$	$c_1 - 10c_2$	
4	$28a_1 - 406a_2$	$-8a_1 + 340a_2$	$-64a_2 + 16b_2 - 4c_2$
	$-7b_1 + 91b_2$	$+2b_1 - 82b_2$	
	$+4c_0 + 2c_1 - 20c_2$	$+20c_2$	
5	$-36a_1 + 810a_2$	$8a_1 - 540a_2$	$80a_2 - 20b_2 + 5c_2$
	$+9b_1 - 189b_2$	$-2b_1 + 132b_2$	
	$+5c_0 - 2c_1$	$c_1 - 32c_2$	

Either via manipulation of formulas (tedious) or interpolation from the first few values we can find the following explicit general form for the polynomials $h_d(z)$ as functions of the indeterminates $a_0, \ldots, c_2$:

$$\begin{aligned} h_d(z) &= (-1)^d \Bigg\{ [-16a_2 + 4b_2 - c_2]\, d\, z^2 + + \left[24a_2 - 6b_2 + \frac{3}{2}c_2\right] d^2\, z \\ &\quad + [-16a_2 + 4b_2 - c_2]\, d\, z + [-8a_1 + 20a_2 + 2b_1 - 2b_2]\, z \\ &\quad + \left[-8a_2 + 2b_2 - \frac{1}{2}c_2\right] d^3 + \left[12a_2 - 3b_2 + \frac{4}{2}c_2\right] d^2 \\ &\quad + \left[8a_1 - 24a_2 - 2b_1 + 3b_2 + (-1)^d c_0 + \frac{1}{2}c_1\right] d \end{aligned}$$

$$+[-4a_1+10a_2+b_1-b_2]\} + \frac{1-(-1)^d}{2}\left[c_1+\frac{1}{2}c_2\right](z-\frac{1}{2})$$

From this we conclude that for $h(d)$ to vanish identically for $d \geq 2$ it is necessary and sufficient to have

$$\begin{aligned} c_2 &= -16a_2+4b_2 \\ c_1 &= -\frac{1}{2}c_2 = 8a_2-2b_2 \\ c_0 &= 0 \\ b_1 &= 4a_1+10a_2+b_2 \end{aligned} \tag{10}$$

and the parameters a_0, a_1, a_2, b_0, b_2 may be chosen independently. (Actually, rank consideration show that the vanishing of $h_2(z)$ and $h_3(z)$ is sufficient to impose these conditions, whereas vanishing of $h_2(z)$ or $h_3(z)$ alone is not).

Let us note that under the mentioned conditions the polynomials $h_{-1}(z)$, $h_0(z)$ and $h_1(z)$ turn out to be:

$$\begin{aligned} h_{-1}(z) &= g_{-1}(z) = a_1 + a_1\,z + a_2\,z^2 \\ h_0(z) &= (-6a_2+b_2)z^2 + (-6a_2+b_2)z - 2a_0 + a_1 + a_2 + b_0 \\ &= (-6a_2)z^2 + (4a_2-4a_1)z - 2a_0 + a_1 - a_2 + g_0(z) \\ h_1(z) &= a_2\,z^2 + (2a_2-a_1)z + a_0 - a_1 + a_2 \\ &= a_2(z+1)^2 - a_1(z+1) + a_0 = g_1(-z-1) \end{aligned} \tag{11}$$

Thus we have shown:

Theorem 5 *For any tridiagonal matrix* $\mathbf{G} = D_\Gamma$ *as specified in (9), the matrix* $\mathbf{H} = D_\Lambda$ *obtained via reduced Legendre transform, i.e.*

$$\mathbf{H} = \mathbf{Q} \cdot \mathbf{G} \cdot \mathbf{Q}^{-1}$$

is again tridiagonal if and only if the conditions in (10) are satisfied.
In this case, the family Λ *is given by (11).*

Note that a *verification* of the if-part of this result is just as easy as the verificational proof of identity (4). Again, we can verify

$$\mathbf{H} \cdot \mathbf{Q} = \mathbf{Q} \cdot \mathbf{G}$$

by comparing the (i,j)-entries of the matrices on both sides with i and j as indeterminates. The same remarks as made at the end of section 3 apply.

The result of theorem 5 can be rephrased in terms of the original Legendre transform. Let

$$\tilde{\mathbf{G}} = \boldsymbol{\Delta}^{-1} \cdot \mathbf{G} \cdot \boldsymbol{\Delta}$$

with $\mathbf{\Delta} = diag\left(\binom{2k}{k}\right)_{k\geq 0}$. Then $\tilde{\mathbf{G}} = D_{\tilde{\Gamma}}$ where $\tilde{\Gamma} = (\tilde{g}_k(z))_{k\in\mathbf{Z}}$ is given by

$$\tilde{g}_k(z) = \begin{cases} \frac{4z+2}{z+1}\, g_{-1}(z) & \text{if } k=-1 \\ g_0(z) & \text{if } k=0 \\ \frac{z}{4z-2}\, g_1(z) & \text{if } k=1 \\ 0 & \text{else} \end{cases}$$

If we impose conditions (10) and furthermore require that $z+1 \,|\, g_{-1}(z)$ (which introduces a further linear constraint: $a_0 = a_1 - a_2$), then we find:

Corollary 6 *Legendre transform*

$$\mathbf{H} = \mathbf{P}\cdot\tilde{\mathbf{G}}\cdot\mathbf{P}^{-1}$$

maps the polynomials

$$\begin{aligned} \tilde{g}_{-1}(z) &= A\,z^2 + B\,z + C \\ \tilde{g}_0(z) &= (A+D)\,z^2 + (-A+B+D)\,z + E \\ \tilde{g}_1(z) &= D\,z^2 \end{aligned}$$

onto

$$\begin{aligned} h_{-1}(z) &= \frac{z+1}{4z+2}\left(A\,z^2 + B\,z + c\right) \\ h_0(z) &= \frac{1}{2}\left[(2D-A)\,z(z+1) + 2E - C\right] \\ h_1(z) &= \frac{z}{4z+2}\left[A\,z^2 + (2A-B)\,z + (A-B+C)\right] \end{aligned}$$

Note that the translation is given by $A = 4a_2$, $B = -2a_2 + 4a_1$, $C = 2(a_1 - a_2)$, $D = c_2/4$ and $E = b_0$.

This is essentially the contents[10] of theorem 1 in [16], where a rather different proof using *creative telescoping* is given. Note that the situation of the Franel-Apéry conjugacy is given by the particular values $A = C = 1, B = 2, D = -8, E = 2$. Schmidt's article contains a number of further examples illustrating this transformation.

5 A brief look at diophantine approximation

In this final section I would like to draw attention to an application of the foregoing to diophantine approximation. The results mentioned here are adapted from [16].

As mentioned in the beginning, Apéry's surprising proof of the irrationality of $\zeta(3)$ is based on the recurrence relation for the Apéry numbers. In a few words, his idea of proof goes as follows (see [20]):

[10] The equivalence is not complete, because the additional linear relation introduced here reduces the number of free parameters by one. But this is a technical matter that can be dealt with.

Consider the two sequences

$$\begin{aligned}(a_n)_{n\geq 0} &= (\,1\,,\,5\,,\,73\,,\,1445\,,\,33001\,,\,\ldots)\\ (b_n)_{n\geq 0} &= (\,0\,,\,6\,,\,\frac{351}{4}\,,\,\frac{62531}{36}\,,\,\frac{11424695}{288}\,,\,\ldots)\end{aligned}$$

satisfying both Apéry's linear recurrence relation

$$n^3\,u_n - (34\,n^3 - 51\,n^2 + 27\,n - 5)\,u_{n-1} + (n-1)^3 u_{n-2} = 0 \tag{12}$$

The surprising fact about $(a_n)_{n\geq 0}$ is that all its terms are *integers*, which is contained in (3), but by no means obvious from (12) alone. The terms of $(b_n)_{n\geq 0}$ are rational numbers, where the denominator of b_n divides $2\,\mathrm{lcm}(1,2,\ldots,n)^3$. This follows from

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

where

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

and this also shows that

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

The joint recurrence (12) for numerator and denominator of this recurrence allows for an estimate of the rate of convergence:

$$\zeta(3) - \frac{b_n}{a_n} = \sum_{k=n+1}^{\infty} \frac{a_k\,b_{k+1} - a_{k+1}\,b_k}{a_k\,a_{k+1}} = O\left(\frac{1}{a_n^2}\right)$$

since

$$a_{n-1}\,b_n - a_n\,b_{n-1} = \frac{a_0\,b_1 - a_1\,b_0}{n^3} = \frac{6}{n^3}$$

The sequences $(p_n)_{n\geq 0}$ and $(q_n)_{n\geq 0}$ with

$$p_n = 2\,\mathrm{lcm}(1,2,\ldots,n)^3\,b_n \quad \text{and} \quad q_n = 2\,\mathrm{lcm}(1,2,\ldots,n)^3\,a_n \quad (n \geq 0)$$

respectively, are *integer* sequences with

$$\lim \frac{p_n}{q_n} = \lim \frac{b_n}{a_n} = \zeta(3)$$

and from an estimate of the growth rate of $(a_n)_{n\geq 0}$ (which is obtained from coefficients of the recurrence (12)) one can show that the convergence of $\left(\frac{p_n}{q_n}\right)_{n\geq 0}$ towards $\zeta(3)$ is rapidly enough (in terms of the growth of the q_n) so as to establish non-rationality of the limit.

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

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

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

As to the Franel recurrence, it has been remarked by Cusick (see a footnote in [20]) that both sequences

$$\begin{aligned} (f_n)_{n\geq 0} &= (1, 2, 10, 56, 346, \ldots) \\ (g_n)_{n\geq 0} &= (0, 3, 12, \frac{208}{3}, \frac{1280}{3}, \ldots) \end{aligned}$$

satisfying the Franel recurrence, which we may write as

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

lead to a diophantine approximation:

$$\lim \frac{g_n}{f_n} = \frac{\pi^2}{8} \tag{13}$$

We have seen above that the equation

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

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

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

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

$$\begin{aligned} (h_n)_{n\geq 0} &= (\, 0\, ,\, 3\, ,\, \frac{93}{8}\, ,\, \frac{1217}{18}\, ,\, \frac{239429}{576}\, ,\, \ldots) \\ (k_n)_{n\geq 0} &= (\, 3\, ,\, -\frac{3}{2}\, ,\, 1\, ,\, -\frac{4}{3}\, ,\, \ldots\, ,\, (-1)^n \frac{3}{n+1}\, ,\, \ldots) \end{aligned}$$

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

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

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

$$\begin{aligned}(c_n)_{n\geq 0} &= \left(0\,,\,6\,,\,90\,,\,\frac{5348}{3}\,,\,\frac{122140}{3}\ldots\right)\\ (d_n)_{n\geq 0} &= (6\,,\,18\,,\,30\,,\,\ldots\,,\,3\,(4n+2)\,,\,\ldots)\end{aligned}$$

As a result, we get the approximations

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

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

References

[1] R. Askey and J. A. Wilson. A recurrence relation generalizing those of Apéry. *J. Austral. Math. Soc. (Series A)*, 36:267–278, 1984.

[2] W. N. Bailey. The generating function for Jacobi polynomials. *J. London Math. Soc.*, 13:243–246, 1938.

[3] P. Cartier. Démonstration "automatique" d'identités et fonctions hypergéometriques [d'après D. Zeilberger]. In *Séminaire Bourbaki*, volume 746, 1991.

[4] L. Comtet. *Advanced Combinatorics*. Reidel, Dordrecht, 1974.

[5] T. Cusick. Recurrences for sums of powers of binomial coefficients. *J. Comb. Theory (A)*, 52:77–83, 1989.

[6] J. Franel. notes published in: *L'Intermédiaire des Mathématiciens*, vol. 1 (1894), 45-47, and vol. 2 (1895), 33-35.

[7] R. W. Gosper. Decision procedure for indefinite hypergeometric summation. *Proc. Nat. Acad. Sci. USA*, 75:40–42, 1978.

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

[9] J. Hornegger. *Hypergeometrische Summation und polynomiale Rekursion.* Diplomarbeit, IMMD-I, Erlangen, 1992.

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

[11] M. A. Perlstadt. Some recurrences for sums of powers of binomial coefficients. *J. Number Theory*, 27:304–309, 1987.

[12] J. Riordan. *Combinatorial Identities.* John Wiley & Sons, Inc., New York-London-Sidney, 1968.

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

[14] A. L. Schmidt. Generalized Legendre polynomials. *J. Reine Angew. Math.*, 404:192–202, 1990.

[15] A. L. Schmidt. Generalized Legendre and q-Legendre polynomials. In *VII Simposium sobre polinomios ortogonales y applicaciones, Granada 1991*, 1991. (to appear in: J. Comp. Appl. Math.).

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

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

[18] V. Strehl. Binomial identities: combinatorial and algorithmic aspects. (in preparation).

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

[20] A. J. van der Poorten. A proof that Euler missed ... Apéry's proof of the irrationality of $\zeta(3)$. *Math. Intelligencer*, 1:195–203, 1978.

[21] H. S. Wilf and D. Zeilberger. Rational functions certify combinatorial identities. *J. Amer. Math. Soc.*, 3:147–158, 1990.

[22] H. S. Wilf and D. Zeilberger. Towards computerized proofs of combinatorial identities. *Bull. Amer. Math. Soc. (New Series)*, pages 77–83, 1990.

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

[24] H. S. Wilf and D. Zeilberger. Rational function certification of multisum/integral/"q" identites. *Bull. Amer. Math. Soc. (New Series)*, 27:148–153, 1992.

[25] D. Zeilberger. A fast algorithm for proving terminating hypergeometric identities. *Discrete Mathematics*, 80:207–211, 1990.

[26] D. Zeilberger. The method of creative telescoping. *J. Symb. Comp.*, 11:195–205, 1991.