A Computer-Algebra-Based Formal Proof of the Irrationality of $\zeta(3)$

Frédéric Chyzak

Joint work with A. Mahboubi, T. Sibut-Pinote, and E. Tassi

May 27, 2014

Apéry's Theorem (1978/1979): The Number $\zeta(3) = \sum_{m=1}^{\infty} \frac{1}{m^3}$ is Irrational

Sketch of proof, as in (van der Poorten, 1979)

- Define:

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

$$v_{n,k} = c_{n,k} u_{n,k}, \quad a_n = \sum_{k=0}^{n} c_{n,k}, \quad b_n = \sum_{k=0}^{n} v_{n,k}.$$

- Prove: (a_n) and (b_n) satisfy the same 2nd-order recurrence, so that

$$0 < \zeta(3) - b_n/a_n = \mathcal{O}\left(a_n^{-2}\right), \qquad a_n = \Theta\left(n^{-3/2}(\sqrt{2}+1)^{4n}\right).$$

- Define $\ell_n = \operatorname{lcm}(1, \dots, n)$ and prove $2\ell_n^3 a_n \in \mathbb{N}$, $2\ell_n^3 b_n \in \mathbb{Z}$.

- Notice $\ell_n = \mathcal{O}(e^n)$ and $e^3(\sqrt{2}+1)^{-4} \simeq 0.59$ to conclude:

$$0 < 2\ell_n^3 \left(a_n \zeta(3) - b_n\right) = \mathcal{O}\left(n^{3/2} e^{3n} (\sqrt{2}+1)^{-4n}\right) \implies \zeta(3) \notin \mathbb{Q}.$$

Apéry's Theorem (1978/1979): The Number $\zeta(3) = \sum_{m=1}^{\infty} \frac{1}{m^3}$ is Irrational

Summary of ingredients of the proof

- Genius to invent the sequences (a_n) and (b_n)
- Elementary number theory
- Deriving same second-order recurrence for (a_n) and (b_n)
- Asymptotic estimates

Apéry's Theorem (1978/1979): The Number $\zeta(3) = \sum_{m=1}^{\infty} \frac{1}{m^3}$ is Irrational

Summary of ingredients of the proof

- Genius to invent the sequences (a_n) and (b_n)
- Elementary number theory
- Deriving same second-order recurrence for (a_n) and (b_n)
- Asymptotic estimates

Focus of the talk on proving the recurrence:

- this part is amenable to computer-algebra methods
- typical use of "creative telescoping" for summation

(Beukers, 1979)

Observe

$$I_n = \ell_n^3 \int_0^1 \int_0^1 \int_0^1 \frac{L_n(x)\, L_n(y)}{1 - u\,(1 - xy)} dx\, dy\, du \in \mathbb{Z} + \mathbb{Z}\,\zeta(3),$$

where

$$L_n(x) = \frac{1}{n!} \frac{d^n}{dx^n} x^n (1 - x)^n \qquad \text{(Legendre orthogonal polynomials)}\,.$$

Integrations by parts and easy bounding yield

$$0 < I_n \leq 2\zeta(3)\, 3^{3n} (\sqrt{2} + 1)^{-4n}.$$

Observing $3^3(\sqrt{2} + 1)^{-4} \simeq 0.79$ implies irrationality.

(Beukers, 1979)

Observe

$$I_n = \ell_n^3 \int_0^1 \int_0^1 \int_0^1 \frac{L_n(x)\, L_n(y)}{1 - u\,(1 - xy)} dx\, dy\, du \in \mathbb{Z} + \mathbb{Z}\,\zeta(3),$$

where

$$L_n(x) = \frac{1}{n!} \frac{d^n}{dx^n} x^n (1-x)^n \qquad \text{(Legendre orthogonal polynomials)}\ .$$

Integrations by parts and easy bounding yield

$$0 < I_n \leq 2\zeta(3)\, 3^{3n} (\sqrt{2} + 1)^{-4n}.$$

Observing $3^3(\sqrt{2}+1)^{-4} \simeq 0.79$ implies irrationality.

Mathematically more elegant, but would not illustrate CA/FP interaction.

Apéry's Recurrence for (a_n) and (b_n)

Second-order recurrence (Apéry, 1978/1979)

$$(n+1)^3 s_{n+1} - (34n^3 + 51n^2 + 27n + 5)\, s_n + n^3 s_{n-1} = 0$$

Cohen and Zagier's "Creative Telescoping" (van der Poorten, 1979)

"[They] cleverly construct

$$q_{n,k} = 4\,(2n+1)\left(k\,(2k+1) - (2n+1)^2\right) c_{n,k}$$

with the motive that

$$(n+1)^3 c_{n+1,k} - (34n^3 + 51n^2 + 27n + 5) c_{n,k} + n^3 c_{n-1,k} = \left[q_{n,j}\right]_{j=k-1}^{j=k}."$$

After summation over k from 0 to $n+1$:

$$(n+1)^3 a_{n+1} - (34n^3 + 51n^2 + 27n + 5)\, a_n + n^3 a_{n-1} = \underbrace{\left[q_{n,j}\right]_{j=-1}^{j=n+1}}_{0-0=0}.$$

Apéry's Recurrence for (a_n) and (b_n)

Second-order recurrence (Apéry, 1978/1979)

$$(n+1)^3 s_{n+1} - (34n^3 + 51n^2 + 27n + 5)\, s_n + n^3 s_{n-1} = 0$$

Cohen and Zagier's "Creative Telescoping" (van der Poorten, 1979)

"[They] cleverly construct

$$Q = 4\,(2n+1)\,\big(k\,(2k+1) - (2n+1)^2\big)$$

with the motive that

$$\Big((n+1)^3 S_n - (34n^3 + 51n^2 + 27n + 5) + n^3 S_n^{-1}\Big) \cdot c = (1 - S_k^{-1})\,(Q \cdot c)\,."$$

After summation over k from 0 to $n+1$:

$$\Big((n+1)^3 S_n - (34n^3 + 51n^2 + 27n + 5) + n^3 S_n^{-1}\Big) \cdot a = \underbrace{[Q \cdot c]_{j=-1}^{j=n+1}}_{0-0=0}.$$

Apéry's Recurrence for (a_n) and (b_n)

Second-order recurrence (Apéry, 1978/1979)

$$(n+1)^3 s_{n+1} - (34n^3 + 51n^2 + 27n + 5)\, s_n + n^3 s_{n-1} = 0$$

Cohen and Zagier's "Creative Telescoping" (van der Poorten, 1979)

"[They] cleverly construct

$$P = (n+1)^3 S_n - (34n^3 + 51n^2 + 27n + 5) + n^3 S_n^{-1}$$

and

$$Q = 4\,(2n+1)\,\left(k\,(2k+1) - (2n+1)^2\right)$$

with the motive that

$$P \cdot c = (1 - S_k^{-1})\,(Q \cdot c)\,."$$

After summation over k from 0 to $n+1$:

$$P \cdot a = [Q \cdot c]_{j=-1}^{j=n+1}.$$

Cohen and Zagier's "Creative Telescoping" (van der Poorten, 1979)

"[They] cleverly construct

$$P = (n+1)^3 S_n - (34n^3 + 51n^2 + 27n + 5) + n^3 S_n^{-1}$$

and

$$Q = 4\,(2n+1)\,\big(k\,(2k+1) - (2n+1)^2\big)$$

with the motive that

$$P \cdot c = (1 - S_k^{-1})\,(Q \cdot c)\,."$$

After summation over k from 0 to $n+1$:

$$P \cdot a = [Q \cdot c]_{j=-1}^{j=n+1}.$$

Skew-polynomial algebras:

$$S_n n = (n+1) S_n, \quad S_k k = (k+1) S_k \quad \text{in} \quad \mathbb{Q}(n,k)\langle S_n, S_k\rangle$$

My Motivations to Reconsider CA from a FP Viewpoint

I do: study computer-algebra algorithms on special functions.

Can an algorithmically-generated encyclopedia be authoritative?

E.g., Dynamic Dictionary of Mathematical Functions (DDMF).

My Motivations to Reconsider CA from a FP Viewpoint

I do: study computer-algebra algorithms on special functions.

Can an algorithmically-generated encyclopedia be authoritative?

E.g., Dynamic Dictionary of Mathematical Functions (DDMF).

Doubts with the litterature related to special-functions algorithms

- some key papers are too informal to assess their correctness / I've lost proofs written too tersely in my own papers
- formal power series vs fractions vs functions? / diagonals, positive parts: Cauchy theorem vs algebraic residues?
- hypergeometric sequence vs hypergeometric term? / holonomic vs rationally holonomic vs D-finite vs ∂-finite vs P-recursive?

My Motivations to Reconsider CA from a FP Viewpoint

I do: study computer-algebra algorithms on special functions.

Can an algorithmically-generated encyclopedia be authoritative?

E.g., Dynamic Dictionary of Mathematical Functions (DDMF).

Doubts with the litterature related to special-functions algorithms

- some key papers are too informal to assess their correctness / I've lost proofs written too tersely in my own papers
- formal power series vs fractions vs functions? / diagonals, positive parts: Cauchy theorem vs algebraic residues?
- hypergeometric sequence vs hypergeometric term? / holonomic vs rationally holonomic vs D-finite vs ∂-finite vs P-recursive?

I want: banish underqualified phrasings and prevent shifts in meaning.
I don't want: reproduce informal interaction with the computer.

Example: Densities of short uniform random walks (Borwein, Straub, Wan, Zudilin, 2012).

Turning our attention to negative integers, we have for $k \geqslant 0$ an integer:

$$W_3(-2k-1) = \frac{4}{\pi^3}\left(\frac{2^k k!}{(2k)!}\right)^2 \int_0^\infty t^{2k} K_0(t)^3 \mathrm{d}t, \tag{78}$$

because the two sides satisfy the same recursion ([BBBG08, (8)]), and agree when $k = 0, 1$ ([BBBG08, (47) and (48)]).

From (78), we experimentally determined a single hypergeometric for $W_3(s)$ at negative odd integers:

Lemma 2. *For $k \geqslant 0$ an integer,*

$$W_3(-2k-1) = \frac{\sqrt{3}\binom{2k}{k}^2}{2^{4k+1}3^{2k}}\, {}_3F_2\left(\left.\begin{matrix}\frac{1}{2}, \frac{1}{2}, \frac{1}{2} \\ k+1, k+1\end{matrix}\right| \frac{1}{4}\right).$$

Proof. It is easy to check that both sides agree at $k = 0, 1$. Therefore we need only to show that they satisfy the same recursion. The recursion for the left-hand side implies a contiguous relation for the right-hand side, which can be verified by extracting the summand and applying Gosper's algorithm ([PWZ06]). $\square$

Summation by Computer Algebra *Is* Used in Proofs

Example: Bounding error in high-precision computation of Euler's constant (Brent, Johansson, 2013).

The "lower" sum L is precisely $\sum_{k=0}^{m/2-1} b_k x^{-2k}$. Replacing k by $2k$ in (21) (as the odd terms vanish by symmetry), we have to prove

$$\sum_{j=0}^{2k} \frac{(-1)^j [(2j)!]^2 [(4k-2j)!]^2}{(j!)^3 [(2k-j)!]^3 32^{2k}} = \frac{[(2k)!]^3}{(k!)^4 8^{2k}}. \tag{23}$$

This can be done algorithmically using the creative telescoping approach of Wilf and Zeilberger. For example, the implementation in the Mathematica package `HolonomicFunctions` by Koutschan [6] can be used. The command

```
a = ((2j)!)^2 / ((j!)^3 32^j);
CreativeTelescoping[(-1)^j a (a /. j -> 2k-j),
    {S[j]-1}, S[k]]
```

outputs the recurrence equation

$$(8+8k)b_{k+1} - \left(1+6k+12k^2+8k^3\right) b_k = 0$$

matching the right-hand side of (23), together with a telescoping certificate. Since the summand in (23) vanishes for $j < 0$ and $j > 2k$, no boundary conditions enter into the telescoping relation, and checking the initial value ($k = 0$) suffices to prove the identity.[1]

[1] Curiously, the built-in `Sum` function in Mathematica 9.0.1 computes a closed form for the sum (23), but returns an answer that is wrong by a factor 2 if the factor $[(4k-2j)!]^2$ in the summand is input as $[(2(2k-j))!]^2$.

Algorithmic theory for Special Functions and Combinatorial Sequences initiated by Zeilberger (1982, 1990, 1991)

- Replace named sequences by linear systems of recurrences
  (+ initial conditions to identify the right solutions)
- Develop algorithms on the level of systems for $+$, $\times$, $\sum$

Implementations exist for Maple, Mathematica, Maxima, etc.

Great success:

- fast evaluation formulae: π, the Catalan constant, ζ-values, β-values
- enumerative combinatorics: heap-ordered trees, q-analogue of totally symmetric plane partitions; positive 3D rook walks; small-step walks
- partition theory: Rogers-Ramanujan and Göllnitz-type identities
- knot theory: colored Jones functions
- mathematical physics: computation of Feynman diagrams

Computer-algebra algorithms apply to Apéry's sums!

- Zeilberger's calculation ($\leq$ 1992) for (a_n)
- Zudilin's alternate proof (1992) by two calls to Zeilberger's algorithm
- Apéry's original calculations using Zeilberger's and Chyzak's algorithms: Salvy's Maple worksheet (2003),
  `http://algo.inria.fr/libraries/autocomb/Apery2-html/apery.html`
- Using difference-field extensions (Schneider, 2007)

Computer-algebra algorithms apply to Apéry's sums!

- Zeilberger's calculation ($\leq$ 1992) for (a_n)
- Zudilin's alternate proof (1992) by two calls to Zeilberger's algorithm
- Apéry's original calculations using Zeilberger's and Chyzak's algorithms: Salvy's Maple worksheet (2003),
  `http://algo.inria.fr/libraries/autocomb/Apery2-html/apery.html`
- Using difference-field extensions (Schneider, 2007)

Computer-algebra algorithms apply to Apéry's sums!

- Zeilberger's calculation ($\leq$ 1992) for (a_n)
- Zudilin's alternate proof (1992) by two calls to Zeilberger's algorithm
- Apéry's original calculations using Zeilberger's and Chyzak's algorithms: Salvy's Maple worksheet (2003),
  `http://algo.inria.fr/libraries/autocomb/Apery2-html/apery.html`
- Using difference-field extensions (Schneider, 2007)

Our formalization follows the Apéry/van der Poorten/Salvy path.

- Fibonacci numbers: $F_{n+2} = F_{n+1} + F_n, \quad F_0 = F_1 = 1.$
- Define (S_n) by: $S_{n+1} = -S_n, \quad S_0 = 1.$
- Introduce $u_n := F_{n+1}^2 + S_n$ and compute the normal forms:

$$\begin{aligned} u_n &= F_{n+1}^2 + S_n, \\ u_{n+1} &= F_n^2 + 2F_nF_{n+1} + F_{n+1}^2 - S_n, \\ u_{n+2} &= F_n^2 + 4F_nF_{n+1} + 4F_{n+1}^2 + S_n, \\ u_{n+3} &= 4F_n^2 + 12F_nF_{n+1} + 9F_{n+1}^2 - S_n. \end{aligned}$$

- Solving a linear system yields: $u_{n+3} - 2u_{n+2} - 2u_{n+1} + u_n = 0.$
- Same process for $v_n := F_nF_{n+2}$ delivers the same recurrence.
- Now, checking initial conditions and induction ends the proof:

$$u_0 = v_0 = 2, \qquad u_1 = v_1 = 3, \qquad u_2 = v_2 = 10.$$

A Generalization: ∂-Finite Sequences (Chyzak, Salvy, 1998)

$$(t_{n,k}) \text{ is } \partial\text{-finite}$$

$$\Updownarrow$$

$$\boxed{\text{the shifts } (t_{n+i,k+j}) \text{ span a finite-dimensional } \mathbb{Q}(n,k)\text{-vector space}}$$

⇒ linear functional equations with rational-function coefficients.

A Generalization: ∂-Finite Sequences (Chyzak, Salvy, 1998)

$(t_{n,k})$ is ∂-finite

$\Updownarrow$

the shifts $(t_{n+i,k+j})$ span a finite-dimensional $\mathbb{Q}(n,k)$-vector space

$\Rightarrow$ linear functional equations with rational-function coefficients.

Examples: Fibonacci numbers; binomial coefficients

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

orthogonal polynomials, Bessel functions.

A Generalization: ∂-Finite Sequences (Chyzak, Salvy, 1998)

$(t_{n,k})$ is ∂-finite

$\Updownarrow$

the shifts $(t_{n+i,k+j})$ span a finite-dimensional $\mathbb{Q}(n,k)$-vector space

$\Rightarrow$ linear functional equations with rational-function coefficients.

Examples: Fibonacci numbers; binomial coefficients

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

orthogonal polynomials, Bessel functions.

Closures under $+$, $\times$, shifts

- Annihilating ideal $\to$ skew Gröbner basis $\to$ normal forms in finite dim.
- Iterative algorithm to search for linear dependencies

$\rightsquigarrow$ simplification and zero test of ∂-finite polynomial expressions.

A Generalization: ∂-Finite Sequences (Chyzak, Salvy, 1998)

$(t_{n,k})$ is ∂-finite

$\Updownarrow$

the shifts $(t_{n+i,k+j})$ span a finite-dimensional $\mathbb{Q}(n,k)$-vector space

$\Rightarrow$ linear functional equations with rational-function coefficients.

Examples: Fibonacci numbers; binomial coefficients

$$\operatorname{ann} \binom{n}{k} = \left\{ L_1 \left(S_n - \frac{n+1}{n+1-k} \right) + L_2 \left(S_k - \frac{n-k}{k+1} \right) : L_1, L_1 \in \mathbb{Q}(n,k)\langle S_n, S_k \rangle \right\};$$

orthogonal polynomials, Bessel functions.

Closures under $+$, $\times$, shifts

- Annihilating ideal $\to$ skew Gröbner basis $\to$ normal forms in finite dim.
- Iterative algorithm to search for linear dependencies

$\rightsquigarrow$ simplification and zero test of ∂-finite polynomial expressions.

A Convoluted Proof of $\sum_{k=0}^{n} \binom{n}{k} = 2^n$

- Define $F_n := \displaystyle\sum_{k=0}^{n} \binom{n}{k}$.
- Prove
$$\binom{n+1}{k} - 2\binom{n}{k} = \left[\frac{-j\binom{n}{j}}{n+1-j}\right]_{j=k}^{j=k+1}$$
as a consequence of
$$\binom{n+1}{k} = \frac{n+1}{n+1-k}\binom{n}{k}, \quad \binom{n}{k+1} = \frac{n-k}{k+1}\binom{n}{k}.$$
- Sum from $k = -1$ to $k = n+1$ to get $F_{n+1} - 2F_n = 0$.
- Now, observing $F_0 = 1$ yields the result.

Algorithms for Summing "Holonomic" ∂-Finite Sequences

Zeilberger's algorithm (1991)
INPUT: a hypergeometric term $f_{n,k}$, that is, first-order recurrences.
OUTPUT: rational functions $p_0(n), \ldots, p_r(n)$, $Q(n,k)$ with minimal r, such that $p_r(n) f_{n+r,k} + \cdots + p_0(n) f_{n,k} = Q(n,k+1) f_{n,k+1} - Q(n,k) f_{n,k}$.

Algorithms for Summing "Holonomic" ∂-Finite Sequences

Zeilberger's algorithm (1991)
Input: a hypergeometric term $f_{n,k}$, that is, first-order recurrences.
Output: rational functions $p_0(n), \dots, p_r(n)$, $Q(n,k)$ with minimal r, such that $p_r(n) f_{n+r,k} + \dots + p_0(n) f_{n,k} = Q(n,k+1) f_{n,k+1} - Q(n,k) f_{n,k}$.

Chyzak's algorithm (2000)
Input: $\begin{cases} \text{a } \partial\text{-finite term } u \text{ w.r.t. } A = \mathbb{Q}(n,k)\langle S_n, S_k\rangle, \\ \text{a Gröbner basis } G \text{ of } \operatorname{ann} u. \end{cases}$
Output: $\begin{cases} P \in \mathbb{Q}(n)\langle S_n\rangle \text{ of minimal possible order}, \\ Q \in A \text{ reduced modulo } G \text{ and such that } P \cdot u = (S_k - 1) Q \cdot u. \end{cases}$

Zeilberger's algorithm (1991)
INPUT: a hypergeometric term $f_{n,k}$, that is, first-order recurrences.
OUTPUT: rational functions $p_0(n), \dots, p_r(n)$, $Q(n,k)$ with minimal r, such that $p_r(n) f_{n+r,k} + \cdots + p_0(n) f_{n,k} = Q(n,k+1) f_{n,k+1} - Q(n,k) f_{n,k}$.

Chyzak's algorithm (2000)

INPUT: $\begin{cases} \text{a } \partial\text{-finite term } u \text{ w.r.t. } A = \mathbb{Q}(n,k)\langle S_n, S_k\rangle, \\ \text{a Gröbner basis } G \text{ of } \operatorname{ann} u. \end{cases}$

OUTPUT: $\begin{cases} P \in \mathbb{Q}(n)\langle S_n\rangle \text{ of minimal possible order}, \\ Q \in A \text{ reduced modulo } G \text{ and such that } P \cdot u = (S_k - 1) Q \cdot u. \end{cases}$

Example: we can get the same 2nd-order operator P for both sides of

$$\underbrace{\sum_{r=0}^{\infty}}_{\text{by C}} \underbrace{\sum_{s=0}^{\infty}}_{\text{by Z}} (-1)^{n+r+s} \binom{n}{r} \binom{n}{s} \binom{n+r}{r} \binom{n+s}{s} \binom{2n-(r+s)}{n} = \underbrace{\sum_{k=0}^{\infty}}_{\text{by Z}} \binom{n}{k}^4.$$

A Skeptic's Approach to Combining FP and CA

"Proving" an algorithm

- would prove all its results satisfy the specifications
- but it is too much work

Instead, use computer-algebra external tool as an oracle

- be as skeptical of the computer algebra as of the human
- approach of choice when checking is simpler than discovering

Inspired by (Harrison, Théry, 1997)

A Program to Derive Recurrences for Apéry's Sums

Concrete sequences …

step	explicit form	operation	input(s)
1	$c_{n,k} = \binom{n}{k}^2 \binom{n+k}{k}^2$	direct	
2	$a_n = \sum_{k=1}^{n} c_{n,k}$	creative telescoping	$c_{n,k}$
3	$d_{n,m} = \frac{(-1)^{m+1}}{2m^3 \binom{n}{m}\binom{n+m}{m}}$	direct	
4	$s_{n,k} = \sum_{m=1}^{k} d_{n,m}$	creative telescoping	$d_{n,m}$
5	$z_n = \sum_{m=1}^{n} \frac{1}{m^3}$	direct	
6	$u_{n,k} = z_n + s_{n,k}$	addition	z_n and $s_{n,k}$
7	$v_{n,k} = c_{n,k} u_{n,k}$	product	$c_{n,k}$ and $u_{n,k}$
8	$b_n = \sum_{k=1}^{n} v_{n,k}$	creative telescoping	$v_{n,k}$

… replaced with abstract analogues: *any* solution of a given GB

step	explicit form	operation	input GB(s)	output GB
1	$c_{n,k} = \binom{n}{k}^2 \binom{n+k}{k}^2$	direct		C
2	$a_n = \sum_{k=1}^{n} c_{n,k}$	creative telescoping	C	A
3	$d_{n,m} = \frac{(-1)^{m+1}}{2m^3 \binom{n}{m} \binom{n+m}{m}}$	direct		D
4	$s_{n,k} = \sum_{m=1}^{k} d_{n,m}$	creative telescoping	D	S
5	$z_n = \sum_{m=1}^{n} \frac{1}{m^3}$	direct		Z
6	$u_{n,k} = z_n + s_{n,k}$	addition	Z and S	U
7	$v_{n,k} = c_{n,k} u_{n,k}$	product	C and U	V
8	$b_n = \sum_{k=1}^{n} v_{n,k}$	creative telescoping	V	B

How Can a Candidate Recurrence be Checked?

Because
$$\binom{n+1}{k} = \frac{n+1}{n+1-k}\binom{n}{k}, \quad \binom{n}{k+1} = \frac{n-k}{k+1}\binom{n}{k},$$
it follows:
$$\binom{n+1}{k} - 2\binom{n}{k} + \left[\frac{j\binom{n}{j}}{n+1-j}\right]_{j=k}^{j=k+1} =$$
$$\binom{n+1}{k} - 2\binom{n}{k} + \frac{(k+1)\binom{n}{k+1}}{n-k} - \frac{k\binom{n}{k}}{n+1-k} =$$
$$\underbrace{\left(\frac{n+1}{n+1-k} - 2 + \frac{k+1}{n-k}\frac{n-k}{k+1} - \frac{k}{n+1-k}\right)}_{=0}\binom{n}{k} = 0.$$

How Can a Candidate Recurrence be Checked?

Because the annihilating (left) ideal I of $\binom{n}{k}$ is generated by the GB

$$g_1 := S_n - \frac{n+1}{n+1-k}, \quad g_2 := S_k - \frac{n-k}{k+1},$$

it follows:

$$S_n - 2 + (S_k - 1)\frac{k}{n+1-k} =$$

$$S_n - 2 + \frac{k+1}{n-k}S_k - \frac{k}{n+1-k} =$$

$$g_1 + \frac{k+1}{n-k}g_2 + \underbrace{\left(\frac{n+1}{n+1-k} - 2 + \frac{k+1}{n-k}\frac{n-k}{k+1} - \frac{k}{n+1-k}\right)}_{=0} \in I.$$

How Can a Candidate Recurrence be Checked?

Because

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

it follows:

$$\binom{n+1}{k} - 2\binom{n}{k} + \left[\frac{j\binom{n}{j}}{n+1-j}\right]_{j=k}^{j=k+1} =$$

$$\binom{n+1}{k} - 2\binom{n}{k} + \frac{(k+1)\binom{n}{k+1}}{n-k} - \frac{k\binom{n}{k}}{n+1-k} =$$

$$\underbrace{\left(\frac{n+1}{n+1-k} - 2 + \frac{k+1}{n-k}\frac{n-k}{k+1} - \frac{k}{n+1-k}\right)}_{=0}\binom{n}{k} = 0.$$

How Can a Candidate Recurrence be Checked?

Because

$$k \neq n+1 \implies \binom{n+1}{k} = \frac{n+1}{n+1-k}\binom{n}{k}, \quad k \neq -1 \implies \binom{n}{k+1} = \frac{n-k}{k+1}\binom{n}{k},$$

it follows:

$$\binom{n+1}{k} - 2\binom{n}{k} + \left[\frac{j\binom{n}{j}}{n+1-j}\right]_{j=k}^{j=k+1} =$$

$$\binom{n+1}{k} - 2\binom{n}{k} + \frac{(k+1)\binom{n}{k+1}}{n-k} - \frac{k\binom{n}{k}}{n+1-k} =$$

$$\underbrace{\left(\frac{n+1}{n+1-k} - 2 + \frac{k+1}{n-k}\frac{n-k}{k+1} - \frac{k}{n+1-k}\right)}_{=0}\binom{n}{k} = 0$$

if $k \neq n+1$, $k \neq n$, and $k \neq -1$.

The Algebraic Disease and a Potential Cure

Explanation:

- ∂-Finite sequences are defined up to values on an algebraic set Δ.
- Closures under $+$, $\times$, S_i are sound, but out of an unknown Δ.
- Meaning of summation is dubious if summation range intersects Δ.

Hope:

- Easy: Discover the recurrences by a Maple session by algorithms.
- Uneasy: Guard each of them by a proviso, but how to get it?

Structure of Our Coq Files

Data of guarded recurrences for each abstracted composite sequence

- human-discovered and -written provisos for each of the recurrences
- Maple-generated coefficients of the recurrences, pretty-printed to Coq
- recurrences written in terms of the proviso name and coefficient names:
  - hypergeometric sequences ($c_{n,k}$, $d_{n,m}$) and indefinite sum (z_n): a GB directly obtained from the explicit form
  - composite under $+$ or $\times$ ($u_{n,k}$ and $v_{n,k}$): a GB directly obtained via algorithmic closure
  - composite under creative telescoping (a_n, $s_{n,k}$, b_n): first, recurrences of the form $P \cdot f = (S_k - 1)Q \cdot f$; then, conversion of the P into a GB

Structure of Our Coq Files

Data of guarded recurrences for each abstracted composite sequence

- human-discovered and -written provisos for each of the recurrences
- Maple-generated coefficients of the recurrences, pretty-printed to Coq
- recurrences written in terms of the proviso name and coefficient names

Proofs of recurrences for each abstracted sequence

- load guarded recurrences for arguments (assumed) and for the composite (being proved)
- assume arguments satisfying relevant recurrences; define the composite as a function of the arguments
- state and prove lemmas (recurrences) for the composite, e.g.:

  Lemma: $\forall c \in \mathbb{Q}^{\mathbb{Z}^2}$, $\forall u \in \mathbb{Q}^{\mathbb{Z}^2}$, $\forall v \in \mathbb{Q}^{\mathbb{Z}^2}$, if c solves C and u solves U and $v = c \times u$, then v solves V.

Structure of Our Coq Files

Data of guarded recurrences for each abstracted composite sequence

- human-discovered and -written provisos for each of the recurrences
- Maple-generated coefficients of the recurrences, pretty-printed to Coq
- recurrences written in terms of the proviso name and coefficient names

Proofs of recurrences for each abstracted sequence

- load guarded recurrences for arguments (assumed) and for the composite (being proved)
- assume arguments satisfying relevant recurrences; define the composite as a function of the arguments
- state and prove lemmas (recurrences) for the composite

Proofs of recurrences for the concrete sequences

- ad-hoc means for initial sequences ($c_{n,k}$, $d_{n,m}$, z_n)
- recurrences for other sequences follows immediately by instantiation
- finally, reduction of fourth-order recurrence for (b_n) to order 2

Sample Creative Telescoping $a_n = \sum_{k=0}^{n} c_{n,k}$

```
Definition precond_rew_Sn (n k : int) := (k != n + 1) /\ (n != -1).
Definition precond_rew_Sk (n k : int) := (k + 1 != 0) /\ (n != 0).
Definition not_D (n k : int) := (n >= 0) && (k >= 0) && (k < n).

Definition rew_Sn_0_0 (n k : int) : rat :=
  let n' : rat := n%:~R in  let k' : rat := k%:~R in
  ((n' + rat_of_Z 1 + k')^2) / ((- n' + - rat_of_Z 1 + k')^2).
Definition rew_Sn (c : int -> int -> rat) := forall (n k : int),
  precond_rew_Sn n k -> c (n + 1) k = rew_Sn_0_0 n k * c n k.
...
Record GB_of_ann c : Type :=
  ann { rew_Sn_ : rew_Sn c; rew_Sk_ : rew_Sk c }.

Variable (c : int -> int -> rat).
Hypothesis (c_ann : GB_of_ann c).

Theorem P_eq_Delta_Q : forall (n k : int), not_D n k ->
  P (c ^~ k) n = Q c n (k + 1) - Q c n k.
Proof.  ...  by field; lia.  Qed.

Let a (n : int) : rat := \sum_(0 <= k < n + 1) (c n k).

Theorem recAperyA (n : int) : n >= 2 -> P a n = 0.
Proof.  rewrite (sound_telescoping P_eq_Delta_Q).  ...  Qed.
```

Sound Creative Telescoping

A lemma instead of a case-by-case analysis

Given $(u_{n,k}) \in \mathbb{Q}^{\mathbb{Z}^2}$, define $U_n = \sum_{k=\alpha}^{n+\beta} u_{n,k}$. Given a set Δ such that

$$(n,k) \notin \Delta \Rightarrow (P \cdot u_{\bullet,k})_n = (Q \cdot u)_{n,k+1} - (Q \cdot u)_{n,k},$$

the following identity holds for any n such that $\alpha \leq n + \beta$:

$$(P \cdot U)_n = \Big((Q \cdot u)_{n,n+\beta+1} - (Q \cdot u)_{n,\alpha}\Big) + \sum_{i=1}^{r} \sum_{j=1}^{i} p_i(n) u_{n+i,n+\beta+j}$$
$$+ \sum_{\alpha \leq k \leq n+\beta \ \wedge \ (n,k) \in \Delta} (P \cdot u_{\bullet,k})_n - (Q \cdot u)_{n,k+1} + (Q \cdot u)_{n,k}.$$

Sound Creative Telescoping

A lemma instead of a case-by-case analysis

Given $(u_{n,k}) \in \mathbb{Q}^{\mathbb{Z}^2}$, define $U_n = \sum_{k=\alpha}^{n+\beta} u_{n,k}$. Given a set Δ such that

$$(n,k) \notin \Delta \Rightarrow (P \cdot u_{\bullet,k})_n = (Q \cdot u)_{n,k+1} - (Q \cdot u)_{n,k},$$

the following identity holds for any n such that $\alpha \leq n + \beta$:

$$(P \cdot U)_n = \Big((Q \cdot u)_{n,n+\beta+1} - (Q \cdot u)_{n,\alpha}\Big) + \sum_{i=1}^{r} \sum_{j=1}^{i} p_i(n) u_{n+i,n+\beta+j}$$
$$+ \sum_{\alpha \leq k \leq n+\beta \,\wedge\, (n,k) \in \Delta} (P \cdot u_{\bullet,k})_n - (Q \cdot u)_{n,k+1} + (Q \cdot u)_{n,k}.$$

In practice: Coq's u, U, P, Q are total maps, extending the mathematical objects.

Sound Creative Telescoping

A lemma instead of a case-by-case analysis

Given $(u_{n,k}) \in \mathbb{Q}^{\mathbb{Z}^2}$, define $U_n = \sum_{k=\alpha}^{n+\beta} u_{n,k}$. Given a set Δ such that

$$(n,k) \notin \Delta \Rightarrow (P \cdot u_{\bullet,k})_n = (Q \cdot u)_{n,k+1} - (Q \cdot u)_{n,k},$$

the following identity holds for any n such that $\alpha \leq n + \beta$:

$$(P \cdot U)_n = \Big((Q \cdot u)_{n,n+\beta+1} - (Q \cdot u)_{n,\alpha} \Big) + \sum_{i=1}^{r} \sum_{j=1}^{i} p_i(n) u_{n+i,n+\beta+j}$$
$$+ \sum_{\alpha \leq k \leq n+\beta \ \wedge \ (n,k) \in \Delta} (P \cdot u_{\bullet,k})_n - (Q \cdot u)_{n,k+1} + (Q \cdot u)_{n,k}.$$

In practice: Coq's u, U, P, Q are total maps, extending the mathematical objects.

Use of the lemma: normalizing the right-hand side (to 0)

- Ill-formed terms should cancel (manual inspection)
- Normalize modulo GB (several copies of stairs: $u_{n,\alpha}$, $u_{n,n+\beta}$)
- Use rational-function normalization to get 0 (Coq's `field`)

Other Parts of the Formalization (Coq + MathComp + CoqEAL)

Elementary number theory

- definition of binomials over $\mathbb{Z}^2$
- standard properties + $1 \leq i \leq j \leq n \implies j\binom{i}{j} \mid \ell_n$

Asymptotic estimates

- of a_n:
  - implicit use of Poincaré–Perron–Kreuser theorem(s) in Apéry's proof
  - replaced with the more elementary $33^n = \mathcal{O}(a^n)$
- of ℓ_n:
  - original proof uses $\ell_n = e^{n+o(1)}$, implied by the Prime Number Theorem
  - replaced with $\ell_n = \mathcal{O}(3^n)$

Numbers: libraries used

- proof-dedicated integers and rationals of MathComp (Gonthier *et al.*)
- computation-dedicated integers and rationals of CoqEAL (Cohen, Mörtberg, Dénès)
- algebraic numbers (Cohen)
- Cauchy reals to encode $\zeta(3)$ as $(z_n)_{n\in\mathbb{N}}$ and a Cauchy-CV proof

Elementary number theory

- definition of binomials over $\mathbb{Z}^2$
- standard properties + $1 \le i \le j \le n \implies j\binom{i}{j} \mid \ell_n$

Asymptotic estimates

- of a_n:
  - implicit use of Poincaré–Perron–Kreuser theorem(s) in Apéry's proof
  - replaced with the more elementary $33^n = \mathcal{O}(a^n)$
- of ℓ_n:
  - original proof uses $\ell_n = e^{n+o(1)}$, implied by the Prime Number Theorem
  - replaced with $\ell_n = \mathcal{O}(3^n)$ [Admitted.]

Numbers: libraries used

- proof-dedicated integers and rationals of MathComp (Gonthier *et al.*)
- computation-dedicated integers and rationals of CoqEAL (Cohen, Mörtberg, Dénès)
- algebraic numbers (Cohen)
- Cauchy reals to encode $\zeta(3)$ as $(z_n)_{n\in\mathbb{N}}$ and a Cauchy-CV proof

End Result (as of May 2014)

We have machine-checked (a stronger statement of):

$$\boxed{\text{Theorem: } \ell_n = \mathcal{O}(3^n) \implies \zeta(3) \notin \mathbb{Q}.}$$

End Result (as of May 2014)

We have machine-checked (a stronger statement of):

$$\boxed{\text{Theorem: } \ell_n = \mathcal{O}(3^n) \implies \zeta(3) \notin \mathbb{Q}.}$$

```
Coq < Print lcmn_asymptotic_bound.
lcmn_asymptotic_bound =
exists (K2 K3 : rat) (N : nat),
  0 < K2  /\  0 < K3  /\  K2 ^ 3 < 33%:~R  /\
  forall (n : nat),
    (N <= n)%N -> (iter_lcmn n)%:~R < K3 * K2 ^ n
      : Prop

Coq < About zeta_3_irrational.
zeta_3_irrational :
lcmn_asymptotic_bound ->
  not (exists (r : rat), (z3 = (r%:CR))%CR)
```

Subjective Conclusions on Getting to Work with Coq (+ MathComp)

An excessively difficult endeavour

- different methodologies over the years $\rightsquigarrow$ documentation out of sync $\rightsquigarrow$ oral transmission
- lack of external documentation $\rightsquigarrow$ read the code?
- no data abstraction
- too difficult to read through notation + coercions + structure inference
- understanding libraries requires a knowledge of Coq's exotic features
- "inverted" learning curve $\rightsquigarrow$ takes $\mathcal{O}(n^2)$ steps instead of $\mathcal{O}(n)$

Subjective Conclusions on Getting to Work with Coq (+ MathComp)

An excessively difficult endeavour

- different methodologies over the years $\rightsquigarrow$ documentation out of sync $\rightsquigarrow$ oral transmission
- lack of external documentation $\rightsquigarrow$ read the code?
- no data abstraction
- too difficult to read through notation + coercions + structure inference
- understanding libraries requires a knowledge of Coq's exotic features
- "inverted" learning curve $\rightsquigarrow$ takes $\mathcal{O}(n^2)$ steps instead of $\mathcal{O}(n)$

Formalization: opposing goals?

- mimicking the mathematical informal interaction
- flushing doubts on proofs/interpretation of mathematical objects

Conclusions and Future Work

- Complete proof by formalizing bound on $\mathrm{lcm}(1, \ldots, n)$
- Test robustness of approach by more examples of sums
- Develop an understanding of why it works, so as to automate our protocol
- Differential analogue: I'm working on proving the second-order ODE for the square-lattice Green function

  $$\int_0^1 \int_0^1 \frac{1}{(1 - xyz)\sqrt{1 - x^2}\sqrt{1 - y^2}} \, \mathrm{d}x \, \mathrm{d}y$$

  using the Coquelicot library
- Dedicated data structure to keep (skew-)polynomials normalized