A Proof that Euler Missed...

Apéry’s Proof of the Irrationality of \( \zeta(3) \)

An Informal Report
Alfred van der Poorten

1. Journées Arithmétiques de Marseille-Luminy, June 1978

The board of programme changes informed us that R.
Apéry (Caen) would speak Thursday, 14.00 “Sur l’irrational-
alité de \( \zeta(3) \).” Though there had been earlier rumours of
his claiming a proof, scepticism was general. The lecture
tended to strengthen this view to rank disbelief. Those who
listened casually, or who were afflicted with being non-
Francophone, appeared to hear only a sequence of unlikely
assertions.

Exercise
Prove the following amazing claims:

1. For all \( a_1, a_2, \ldots \)

\[
\sum_{k=1}^{\infty} \frac{a_1 a_2 \cdots a_{k-1}}{(x + a_1) \cdots (x + a_k)} = \frac{1}{x}.
\]

2. \( \zeta(3) = \sum_{n=1}^{\infty} \frac{1}{n^3} = \frac{\pi^2}{6} = \sum_{n=1}^{\infty} \frac{1}{n^2 \binom{2n}{n}} \). (1)

3. Consider the recursion:

\[
n^2 u_n - (n - 1)^2 u_{n-2} = (11n^2 - 11n + 3) u_{n-1},
\]

\[n \geq 2,\ (4)\]

Let \( \{b_n\} \) be the sequence defined by \( b_0 = 3, b_1 = 3 \) and
the recursion; then the \( b_n \) are integers! Let \( \{a_n\} \) be the
sequence defined by \( a_0 = 0, a_1 = 5 \) and the recursion; then
the \( a_n \) are rational numbers with denominator dividing
\([1, 2, \ldots, n]^2\).

4. \( a_n/b_n \rightarrow \zeta(3) = \pi^2/6; \) indeed the convergence is so fast
as to imply that for all integers \( p, q \) with \( q \) sufficiently large
relative to \( \epsilon > 0 \)

\[|n^2 - \frac{p}{q}| > \frac{1}{q^{\theta + \epsilon}}, \ \theta' = 11.85078 \ldots .\]

I heard with some incredulity that, for one, Henri
Cohen (Bordeaux, now Grenoble) believed that these claims
might well be valid. Very much intrigued, I joined Hendrik
Lenstra (Amsterdam) and Cohen in an evening’s discussion
in which Cohen explained and demonstrated most of the
details of the proof. We came away convinced that Profes-
sseur Apéry had indeed found a quite miraculous and
magnificent demonstration of the irrationality of \( \zeta(3) \). But we
remained unable to prove a critical step.

2. For the Nonexpert Reader

A number \( \beta \) is irrational if it is not of the form \( p_0/q_0; p_0, q_0 \)
integers (\( \in \mathbb{Z} \)). A rational number \( b \) is characterised by the
property that for \( p, q \in \mathbb{Z} \) (\( q > 0 \)) and \( b \neq p/q \) there exists
an integer \( q_0 > 0 \), of course) such that

\[|b - \frac{p}{q}| > \frac{1}{q^{\theta + \epsilon}}.\]
On the other hand for irrational $\beta$ there are always infinitely many $p/q$ (for instance, the convergents of the continued fraction expansion of $\beta$) such that

$$|\beta - \frac{p}{q}| < \frac{1}{q^2}.$$  

Plainly this yields a criterion for irrationality.

If there is a $\delta > 0$ and a sequence \( \{p_n/q_n\} \) of rational numbers such that $p_n/q_n \neq \beta$ and

$$|\beta - \frac{p_n}{q_n}| < \frac{1}{q_n^{1+\delta}} \quad n = 1, 2, \ldots$$

then $\beta$ is irrational.

A successful application of the criterion may yield a measure of irrationality:

If $|\beta - (p_n/q_n)| < 1/q_n^{1+\delta}$, and the $q_n$ are monotonic increasing with $q_n < q_{n+1}^{1+\delta}$ (for $n$ sufficiently large relative to $\kappa > 0$), then for all integers $p, q > 0$ sufficiently large relative to $\epsilon > 0$, $|\beta - (p/q)| > 1/(q^{1+\delta}(\delta - \kappa)\epsilon)$.

For example if the sequence $\{q_n\}$ increases geometrically we may take $\kappa > 0$ arbitrarily small so that $1 + (1/\delta)$ becomes an irrationality degree for $\beta$. To see the claim suppose that $|\beta - (p/q)| < 1/q^\tau$ and select $n$ so that $q_{n-1}^{1+\delta} < q_n^{1+\delta}$. Then

$$\frac{1}{q_n} \leq |\beta - p/q_n| = |\beta - P_n/q_n| + |\beta - P_n/q_n| < \frac{1}{q_n^{1+\delta}} + \frac{1}{q_n^{1+\delta}} < \frac{2}{q_n^{1+\delta}}.$$ 

Hence $\frac{1}{\epsilon} < q_n < q_n^{1+\delta} = q_n^{1+\delta}/(1+\delta)$ or $\tau < (1+\delta)/(\delta - \kappa) + \epsilon$ as claimed. This argument is effective (the “sufficiently large” requirements can be made explicit).

It is well-known (the theorem of Thue–Siegel–Roth) that for $\beta$ algebraic (a zero of a polynomial $a_0x^n + a_1x^{n-1} + \ldots + a_n, a_i \in \mathbb{Z}$) always: $|\beta - (p/q)| > 1/q^{2+\epsilon},$ for $q$ sufficiently large relative to $\epsilon > 0$. So if $\beta$ is too well approximable by rationals ($\delta > 1$ above) then $\beta$ is not algebraic, but transcendental. Unfortunately only a set of measure zero of transcendental numbers can be detected in this way, whilst, since the set of algebraic numbers is countable, almost all numbers are transcendental.

It is notoriously difficult to prove that any given naturally occurring number is irrational, let alone transcendental. One may be fortunate: for example the usual series for $e$ implies immediately (easy exercise) that $e$ is irrational. In the case of the (Riemann) $\zeta$-function: $\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$ (Re $s > 1$) there is the quite well-known fact that

$$\zeta(2k) = \sum_{n=1}^{\infty} n^{-2k} = \left(-1\right)^{k-1} \frac{(2\pi)^{2k}}{2 \cdot (2k)!} B_{2k},$$

$k = 1, 2, \ldots$ (5)

where the Bernoulli numbers, $B_m$, are rational ($\zeta(2) = \pi^2/6$, $\zeta(4) = \pi^4/90, \zeta(6) = \pi^6/945, \ldots$). There are some classical techniques$^1$ for detecting the irrationality of powers of $\pi$, but it is most useful to appeal to the theorem of Hermite–Lindemann (whereby $e^{\alpha}$ is transcendental for algebraic $\alpha \neq 0$) whence $\pi$ is transcendental (because $e^{\pi i} = -1$) and so a fortiori its powers are irrational. So it has long been known that $\zeta(2k)$ is irrational, $k = 1, 2, \ldots$ On the other hand there are no useful analogous closed evaluations of $\zeta$ at odd arguments.$^2$ Incidentally, (5) is demonstrated quite easily.

The Bernoulli numbers are defined by the generating function (a nontrivial example of an even function!)

$$\frac{z}{e^z - 1} = \frac{1}{2} + \frac{1}{2} \sum_{m=0}^{\infty} \frac{B_{2m}}{(2m)!} z^{2m},$$

hence by the recursion

$$(n+1)B_0 + (n+1)B_1 + \ldots + (n+1)B_n = 0,$$

$B_0 = 1, \quad B_1 = -\frac{1}{2}, \quad n = 2, 3, \ldots$$

On the other hand it is well-known that

$$\sin \pi z = \pi z \prod_{n=1}^{\infty} \left(1 - \frac{z^2}{n^2}\right).$$


$^2$ There is however a famous formula of Ramanujan: let $\alpha$ and $\beta$ be positive numbers such that $\alpha \beta = \pi^2$. Then if $n$ is any positive integer

$$\frac{1}{\alpha n} \left(\frac{1}{2} \zeta(2n + 1) + \sum_{k=1}^{\infty} \frac{k^{2n+1} - 1}{e^{2\alpha k} - 1}\right) = (-\beta)^{-n} \left(\frac{1}{2} \zeta(2n + 1) + \sum_{k=1}^{\infty} \frac{k^{2n+1} - 1}{e^{2\beta k} - 1}\right) - 2 \sum_{k=0}^{n+1} (-1)^k B_{2k} \frac{B_{2n+2-2k}}{(2k)! (2n + 2 - 2k)!} \alpha^{n+1-k} \beta^k.$$

Taking $\alpha$ a rational multiple of $\pi$ one sees that $\zeta(2n + 1)$ is given as a rational multiple of $\pi^{2n+1}$ plus two very rapidly convergent series. See for example: Bruce C. Berndt: Modular transformations and generalizations of several formulae of Ramanujan: *Rocky Mountain J. of Maths.* 7 (1977) 147–189. Indeed the above formula is the natural analogue of Euler’s formula (5). The cited paper gives many other formulas and detailed references.
\[
\frac{\pi \cos \pi z}{\sin \pi z} = \pi \cot \pi z = 1 - \sum_{n=1}^{\infty} \left( \frac{2z}{n^2 - z^2} \right).
\]

But

\[
\pi z \cot \pi z = \pi iz \frac{e^{iz} + e^{-iz}}{e^{iz} - e^{-iz}} = \frac{2\pi iz}{e^{2\pi iz} - 1} + \pi iz = \sum_{m=0}^{\infty} (-1)^m \frac{(2\pi)^{2m}}{(2m)!} B_{2m} z^{2m},
\]

and on the other hand

\[
\pi z \cot \pi z = 1 = 2 \sum_{m=1}^{\infty} \left( \sum_{n=1}^{\infty} \frac{1}{n^{2m}} \right) z^{2m}.
\]

Comparing coefficients one has (5). With a little ingenuity one can avoid a direct appeal to the infinite product for \(\sin \pi z\) or to the expansion for \(\pi \cot \pi z\).

Indeed, proving the irrationality of \(\xi(2n+1)\), \(n = 1, 2, \ldots\) constitutes one of the outstanding problems of the theory (ranking with the arithmetic nature of \(\gamma = \lim_{n \to \infty} (1 + \ldots + 1/n) - \log n\), and of \(\pi, e, \pi + e, \ldots\) which are yet undetermined). It is some measure of Apéry's achievement that these questions have been considered by mathematicians of the top rank over the past few centuries without much success being achieved.

3. Some Irrelevant Explanations

For much of the following details I am indebted to Henri Cohen. All this due to Apéry, of course. The identity

\[
\sum_{k=1}^{K} \frac{a_1 a_2 \ldots a_{k-1}}{(x + a_1) \ldots (x + a_k)} = \frac{1}{x} \frac{a_1 a_2 \ldots a_K}{x(x + a_1) \ldots (x + a_K)}
\]

follows easily on writing the right-hand side as \(A_0 - A_K\) and noting that each term on the left is \(A_{k-1} = a_k\). This explains (1). Now put \(x = n^2, a_k = -k^2\), and take \(k \leq K < n - 1\), to obtain

\[
\sum_{k=1}^{n-1} \frac{(-1)^{k-1}(k-1)!^2}{(n^2 - 1)^2 \ldots (n^2 - k^2)} = \frac{1}{n^2 - 2} - \frac{(-1)^{n-1}(n-1)!^2}{(n^2 - n^2 - 1^2) \ldots (n^2 - (n-1)^2)} = \frac{1}{n^2 - 2} - \frac{2(-1)^{n-1}}{n^2 (2n/3)}.
\]

Writing

\[
\epsilon_{n,k} = \frac{1}{2} \frac{k!^2(n-k)!}{k^3(n+k)!},
\]

because

\[
(-1)^k n (\epsilon_n,k - \epsilon_{n-1,k}) = \frac{(-1)^{k-1}(k-1)!^2}{(n^2 - 1^2) \ldots (n^2 - k^2)}
\]

we have

\[
\sum_{n=1}^{N} \sum_{k=1}^{n-1} (-1)^k (\epsilon_n,k - \epsilon_{n-1,k}) = 0
\]

\[
= \sum_{n=1}^{N} \frac{1}{n^3} - 2 \sum_{n=1}^{N} \frac{(-1)^{n-1}}{n^3 (2n/3)} = \sum_{k=1}^{N} (-1)^k \epsilon_{n,k} - \epsilon_{n,k}
\]

\[
= \sum_{k=1}^{N} \frac{(-1)^k}{2k^3} \frac{1}{k} \left( \frac{N+k}{N} \right) + \frac{1}{2} \sum_{k=1}^{N} (-1)^{k-1} \frac{1}{k^3} \frac{1}{k}
\]

and on noting that as \(N \to \infty\) the first term on the right vanishes, we have (2).

4. Some Nearly Relevant Explanations

All this is quite irrelevant to the proof. It would suffice to introduce the quantities \(k \leq n\)

4 Actually the formula (2) is quite well known: it was observed some years ago by Raymond Ayoub (Penn State) and it in fact appears in print: Margrethe Munthe Hjortnaes: Überführung av rekken \(\sum_{k=1}^{\infty} (1/k^3)\) til et bestemt integral Proc. 12th Cong. Scand. Math., Lund 10–15 Aug. 1953 (Lund 1954); independently again it was noticed by R. William Gosper, Jr. (Palo Alto), see Gosper's paper: A calculus of series rearrangements Algorithms and Complexity, New Directions and Recent Results, ed. J. Traub (Academic Press, 1976) 121–151, for relevant techniques. Henri Cohen remarked that the formula is:

\[
\xi(3) = \frac{5}{4} \log^2 \omega + \frac{2 \pi^2}{15} \log \omega - \frac{2}{3} \log^3 \omega
\]

(with \(\omega = \frac{1}{2}(1 + \sqrt{5})\) and \(\xi(x) = \sum x^3/n^3\), the trilogarithm). Hjortnaes and Ayoub, and respectively Gosper note the integral representations (easily shown equivalent)

\[
\xi(3) = \int_{0}^{1} \log^2 t \coth t dt
\]

\[
\xi(3) = \int_{0}^{1/2} \arcsinh t^2 dt
\]

In the case \(\xi(2)\) the formula is even better known. It is, for example, referred to by Z. R. Melzak: Introduction to Concrete Mathematics (Wiley, 1973), p. 85 (but the suggested proof is not quite appropriate). It may be proved by slightly varying the argument in Section 3 – multiply by \((-1)^{n-1}\) instead of dividing by \(n\). Many formulas similar to (2) and (2) appear in the literature and the folklore.
\[ c_{n,k} = \sum_{m=1}^{n} \frac{1}{m^3} + \sum_{m=1}^{k} \frac{(-1)^{m-1}}{2m^3(m)(n+m)}, \]

and to remark that plainly \( c_{n,k} \to \xi(3) \) as \( n \to \infty \), uniformly in \( k \). One might hope that a sequence \( c_{n,k} \) already implies the irrationality of \( \xi(3) \) (say, the diagonal, with \( k = n \)) but this is not quite so. To see this, it is useful to prove a lemma:

**Lemma.**

\[ 2c_{n,k}(n+k) \in \mathbb{Z} + \frac{\mathbb{Z}}{2^3} + \ldots + \frac{\mathbb{Z}}{n^3} = \frac{\mathbb{Z}}{[1, 2, \ldots, n]^3} \]

(equivalently: \( 2[1, 2, \ldots, n]^3 c_{n,k}(n+k) \) is an integer).

**Proof.** We check the number of times that any given prime \( p \) divides the denominator. But

\[ \binom{n+k}{m} / \binom{n+m}{m} = \binom{n+k}{k-m} / \binom{k}{m} \]

so, because

\[ \text{ord}_p(\binom{n}{m}) = \left\lfloor \frac{\log n}{\log p} \right\rfloor - \text{ord}_p m = \text{ord}_p [1, \ldots, n] - \text{ord}_p m, \]

we have

\[ \text{ord}_p(3m^3(m)(n+k))/(n+k) = \text{ord}_p(3m^3(m)(k-k))/(n+k) \]

\[ \leq 3(\text{ord}_p m) + \left\lfloor \frac{\log n}{\log p} \right\rfloor + \left\lfloor \frac{\log k}{\log p} \right\rfloor - 2(\text{ord}_p m), \]

which yields the assertion, because \( m \leq k \leq n \). \( \square \)

We remark that those who know it well\(^5\) know that for \( n \) sufficiently large relative to \( \varepsilon > 0 \),

\[ [1, 2, \ldots, n] \leq e^{n(1+\varepsilon)} \]

(roughly: \([1, 2, \ldots, n] = \Pi_{p \leq n}^{\log \log p} \leq \Pi_{p \leq n}^{\log n} \approx n^{\log \log n} = e^n\)). It will turn out that the \( c_{n,k} \) have too large a denominator relative to their closeness to \( \xi(3) \). Hence to apply the irrationality criterion we must somehow accelerate the convergence. Apéry described this process as follows:

Consider two triangular arrays (defined for \( k \leq n \)) with entries \( d_{n,k}^{(0)} = c_{n,k}(n+k) \) and \( (n+k) \) respectively. We recall

\[ 5 \text{ Those who know it really well write } \log[1, \ldots, n] = \sum_{m=1}^{n} \theta(n^{1/m}) = \psi(n) \text{ where } \theta(n) = \sum_{m \leq n} \log m. \text{ Then it is known that } \psi(n)/n \leq 1.03883 \ldots \text{ (with maximum at } n = 113) \text{ and indeed } \psi(n) - n < (0.0242334 \ldots) n/\log n \text{ for } n \geq 525752; \text{ See J. Barkley Rosser and Lowell Schoenfeld: Math. Comp. 29 (1975), 243–269.} \]
\[ n^3 a_n - P(n-1)a_{n-1} + (n-1)^3 a_{n-2} = 0, \]
\[ n^3 b_n - P(n-1)b_{n-1} + (n-1)^3 b_{n-2} = 0, \]

one multiplies the first equation by \( b_{n-1} \), the second by \( a_{n-1} \), to obtain

\[ n^3 (a_n b_{n-1} - a_{n-1} b_n) = (n-1)^3 (a_{n-1} b_{n-2} - a_{n-2} b_{n-1}). \]

Recalling \( a_1 b_0 - a_0 b_1 = 6 \times 1 - 0 \times 5 = 6 \), this cleverly yields

\[ a_n b_{n-1} - a_{n-1} b_n = \frac{6}{n^3}. \]  \( \text{(7)} \)

Seeing that \( \xi(3) - a_0 b_0 = \xi(3) \), it is easily induced\(^6\) that

\[ |\xi(3) - \frac{a_n}{b_n}| = \frac{6}{k^{n+1} k^3 b_k b_{k-1}}. \]

so

\[ \xi(3) - \frac{a_n}{b_n} = O(b_n^{-2}). \]

On the other hand the recurrence relation makes it easy to estimate \( b_n \), at any rate asymptotically. We have

\[ b_n = (34 - 51n^{-1} + 27n^{-2} - 5n^{-3}) b_{n-1} + (1 - 3n^{-1} + 3n^{-2} - n^{-3}) b_{n-2} = 0, \]

and since the polynomial \( x^2 - 34x + 1 \) has zeroes \( 17 \pm 12 \sqrt{2} = (1 \pm \sqrt{2})^4 \) we readily conclude that \( b_n = O(\alpha^n) \),

\[ \alpha = (1 + \sqrt{2})^4. \]

In fact Cohen has, more precisely, calculated that

\[ b_n = \frac{(1 + \sqrt{2})^2}{(2\sqrt{2})^{3/2}} \frac{(1 - 48 - 15\sqrt{2})}{n^{3/2}} + O(n^{-2}). \]  \( \text{(8)} \)

We have to recall that the \( a_n \) are not integers. But writing

\[ p_n = 2[1, 2, \ldots, n]^3 a_n, q_n = 2[1, 2, \ldots, n]^3 b_n \]

we have \( p_n, q_n \in \mathbb{Z} \) and

\[ a_n = O(\alpha^n e^{3n}), \quad \xi(3) - \frac{p_n}{q_n} = O(\alpha^{-2n}) = O(q_n^{-1+\delta}), \]

\[ \delta = \frac{\log \alpha - 3}{\log \alpha + 3} = 0.080529 \ldots > 0. \]

Hence, by the irrationality criterion, \( \xi(3) \) is indeed irrational, and moreover, because \( \delta^{-1} = 12.417820 \ldots \) we have:

For all integers \( p, q > 0 \) sufficiently large relative to \( \varepsilon > 0 \):

\[ |\xi(3) - \frac{p}{q}| > q^{-\theta + \varepsilon}, \quad \theta = 13.417820 \ldots \]

6. Some Trivial Verifications

To convince ourselves of the validity of Apéry’s proof we need only complete the following exercise.

Exercise

Prove the following identities:

\( \mathbf{5} \) Let

\[ b_n = \sum_{k=0}^{n} \binom{n}{k}^2 \binom{k}{n-k}^2, \quad a_n = \sum_{k=0}^{n} \binom{n}{k}^2 \binom{n+k}{k} c_{n,k}, \]

\[ c_{n,k} = \frac{1}{m^2} \sum_{m=1}^{n} \frac{(-1)^{m-1}}{2m^2 \binom{n}{m} \binom{n+m}{m}}. \]

Then \( a_0 = 0, a_1 = 6; b_0 = 1, b_1 = 5 \) and each sequence \( \{a_n\} \)

and \( \{b_n\} \) satisfies the recurrence (2).

In the same spirit, the case of \( \xi(2) \) requires:

\( \mathbf{5} \) Let

\[ b'_n = \sum_{k=0}^{n} \binom{n}{k}^2 \binom{n+k}{k}, \quad a'_n = \sum_{k=0}^{n} \binom{n}{k}^2 \binom{n+k}{k} c'_{n,k}, \]

\[ c'_{n,k} = 2 \sum_{m=1}^{n} \frac{(-1)^{m-1}}{m^2} + \sum_{m=1}^{n} \frac{(-1)^{m+k-1}}{m^2 \binom{n}{m} \binom{n+m}{m}}. \]

Then \( a'_0 = 0, a'_1 = 5; b'_0 = 1, b'_1 = 3 \) and each sequence \( \{a'_n\} \)

and \( \{b'_n\} \) satisfies the recurrence (4).

It is useful to notice that very little more than just proving these claims is required for Apéry’s proof. After all, it is quite plain that \( a_n/b_n \rightarrow \xi(3); \) the \( b_n \) are integers, and the lemma of Section 4 shows that the \( a_n \) are “near-integers.”

In Section 5 we showed that given that the sequences satisfy the recursion (2) the irrationality of \( \xi(3) \) follows because

\(^6\) Write \( \xi(3) - \frac{a_n}{b_n} = \chi_n \) and note that we have \( \chi_n - \chi_{n-1} = \frac{6}{n^3 b_n b_{n-1}} \) and \( \chi_0 = 0. \)
from $\log \alpha > 3 (\alpha = (1 + \sqrt{2})^4)$ we obtain $\delta > 0$. Thus, as implied in various asides, most of the earlier argument is quite irrelevant. Indeed I am indebted to John Conway for the remark that even $\delta$ is irrelevant.

Exercise

Be the first in your block to prove by a 2-line argument that $\zeta(3)$ is irrational.\(^7\)

6. Given the definitions of $\delta$ show that $a_n b_{n-1} - a_{n-1} b_n = b_n^{-3}$ and $b_n = O(\alpha^n)$ with $\alpha = (1 + \sqrt{2})^4$. Conclude that $\zeta(3)$ is irrational because $\log \alpha > 3$.

Exercise

Astound your friends with an excellent irrationality measure for $\pi^2$.

6. Given the definitions of $\delta$ show that $a_n b_{n-1} - a_{n-1} b_n = 5(-1)^{n+1} n^{-2}$ and $b_n = O(\alpha^n)$ with $\alpha = (1/2)(1 + \sqrt{5})^5$. Conclude that $|\pi^2 - (p/q)| > q^{-(9+\epsilon)}$ with $\theta = 11.850782 \ldots$ for all integers $p, q > Q(\epsilon)$.

7. ICM '78, Helsinki, August 1978

Neither Cohen nor I had been able to prove $\delta$ in the intervening 2 months. After a few days of fruitless effort the specific problem was mentioned to Don Zagier (Bonn), and with irrigating speed he showed that indeed the sequence $\{b_n\}$ satisfies the recurrence (4). This more or less broke the dam and $\delta$ and $\delta$ were quickly conquered. Henri Cohen addressed a very well-attended meeting at 17.00 on Friday, August 18 in the language of the majority, proving $\delta$ and explaining how this implied the irrationality of $\zeta(3)$. Apéry then made some remarks on the status of the French language, and alluded to the underlying motivation (as mentioned in Section 3) for his astonishing proof.

Exercise

A red herring?

7. Show that

$$\zeta(3) = \frac{6}{117} - \frac{64}{535} - \frac{729}{1436} - \frac{4096}{3105} - \frac{n^6}{34n^3 + 51n^2 + 27n + 5} \ldots$$

and deduce that $\zeta(3) = 1.202056903 \ldots$ is irrational.

7. Show that

$$\zeta(2) = \frac{\pi^2}{6} = \frac{5}{3} + \frac{1}{25+16} + \frac{1}{69+81} + \frac{1}{135+256} + \frac{n^4}{223+\ldots}$$

and deduce that $\pi^2$ has irrationality degree at most $11.850782 \ldots$.

8. Some Rather Complicated but Ingenious Explanations

According to a dictum of Littlewood any identity, once verified, is trivial. Surely $\delta$ is very nearly a counterexample. The following is principally due to Zagier and Cohen. Incidentally, we first considered $\delta$ which appeared simpler, but this was because we had failed to notice that

$$\sum_{k=0}^{n} \sum_{l=0}^{k} \binom{n}{k} \binom{k}{l} \binom{2n-l}{2n-k}^2 = \sum_{k=0}^{n} \binom{n}{k}^2 \binom{2n-k}{2n-k}^2.$$ 

Now writing $n - k$ for $k$ links the arrays of Section 4 to $\delta$. 

7. The author does not pretend to be able to do this. Notice that in fact even less is needed: it is sufficient to show $a_n b_{n-1} - a_{n-1} b_n = O(\gamma^n)$ and $b_n = O(\beta^n)$, with $\log \beta - \log \gamma > 3$.

8. Though we have long known that $\zeta(2)$ is irrational, Apéry's result in this case is significant. The irrationality degree for $\pi^2$ is the best known; the irrationality degree implied for $\pi$ is 23.701564. . . . These results compare very favourably with those of Mahler: $|\pi - (p/q)| > q^{-42}$: On the approximation of $\pi$ Proc. K. Ned. Akad. Wet. Amsterdam A, 56 (= Indag. Math. 15) (1953) 29-42, and an indication that $|\pi - (p/q)| > q^{-30}$; see also K. Mahler: Applications of some formulae by Hermite to the approximation of exponentials and logarithms. Math. Annalen 168 (1967) 200-227. Wirsing announced $|\pi - (p/q)| > q^{-71}$ and Mignotte proved that (for $q$ sufficiently large) $|\pi - (p/q)| > q^{-20}$; this is the best known result. It should be noted that the cited result depends on deep techniques and complicated estimations in transcendence theory as contrasted with the essentially elementary methods in Apéry's proof.

Mignotte (op. cit.) also shows that $|\pi^2 - (p/q)| > q^{-18}$, which is weaker than Apéry's result.
It is quite convenient to write:

\[ b_{nk} = \binom{n}{k}^2 \binom{n+k}{k}^2, \quad a_{nk} = b_{nk} c_{nk}, \]

\[ b_n = \sum_{k=0}^{n} b_{nk}, \quad a_n = \sum_{k=0}^{n} b_{nk} c_{nk}. \]

Then we wish to show that

\[ \sum_{k} (n+1)^3 b_{n+1,k} - (34n^3 + 51n^2 + 27n + 5) b_{nk} + n^3 b_{n-1,k} = 0. \]

We cleverly construct

\[ B_{nk} = 4(2n+1)(k(2k+1) - (2n+1)^2) \binom{n}{k}^2 \binom{n+k}{k}^2, \]

with the motive that

\[ B_{nk} - B_{nk-1} = (n+1)^3 \binom{n+1}{k}^2 \binom{n+1+k}{k}^2 - \]

\[ - (34n^3 + 51n^2 + 27n + 5) \binom{n}{k}^2 \binom{n+k}{k}^2 + \]

\[ + n^3 \binom{n-1}{k}^2 \binom{n-1+k}{k}^2, \]

and, \textit{0 mirabile dictu}, the sequence \{b_n\} does indeed satisfy the recurrence (2) by virtue of the method of creative telescoping (by the usual conventions: \(B_{nk} = 0\) for \(k < 0\) or \(k > n\); note also that \(P(n) = 34n^3 + 51n^2 + 27n + 5\) implies \(P(n-1) = -P(-n)\).)

The rest is plain sailing (or is it plane sailing?). We notice that

\[ (n+1)^3 c_{n+1,k} c_{n+1,k} - P(n) b_{nk} c_{nk} + n^3 b_{n-1,k} c_{n-1,k} = \]

\[ = (B_{nk} - B_{nk-1}) c_{nk} + \]

\[ + (n+1)^3 b_{n+1,k} (c_{n+1,k} - c_{nk}) - \]

\[ - n^3 b_{n-1,k} (c_{nk} - c_{n-1,k}). \]  

(9)

Clearly

\[ c_{nk} - c_{n-1,k} = \frac{1}{n^3} \sum_{m=1}^{k} (-1)^m (m-1)!^2 (n-m-1)! \]

\[ = \frac{1}{n^3} \sum_{m=1}^{k} \left( \frac{(-1)^m m!^2 (n-m-k)!}{n^2 (n+m)!} \right) - \]

\[ - \frac{(-1)^m (m-1)!^2}{n^2 (n+m+1)!}. \]

\[ = \frac{(-1)^k k!^2 (n-k-1)!}{n^2 (n+k)!}, \]

whilst not even a minor miracle is required to write down

\[ c_{n,k} - c_{n-1,k}. \] After some massive reorganisation (9) becomes

\[ A_{nk} - A_{n-1,k} \quad \text{with} \quad A_{nk} = B_{nk} c_{nk} + \]

\[ + \frac{5(2n+1)(-1)^{k-1}k}{n(n+1)} \binom{n}{k}^2 \binom{n+k}{k}, \]

and we have completed (5), and, in passing, proved (3). This of course verifies Apéry's claim to have proved \(3(3)\) irrational.

9. The Case of \(\xi(2)\)

The arguments required to deal with the exercises (2) are quite similar to those already described. It may however be a kindness to the reader to reveal that it would be wise to take

\[ B_{nk} = (k^2 + 3(2n+1)k - 11n^2 - 9n - 2) \binom{n}{k}^2 \binom{n+k}{k}, \]

\[ A_{nk} = B_{nk} c_{nk} + 3(-1)^{n+k-1} \frac{(n-1)!}{(k-1)!}. \]

Moreover

\[ c_{nk} - c_{n-1,k} = 2(-1)^{n+k-1} \frac{k!^2 (n-k-1)!}{m(n+k)!}, \]

and

\[ b_n = \frac{\left(\frac{1}{2}(1+\sqrt{5})\right)^4}{2\pi\sqrt{5} + 2\sqrt{5}} \frac{(\frac{1}{2}(1+\sqrt{5}))^5n}{n} \left(1 + O(n^{-1})\right); \]  

(10)

(also note that if \(Q(n) = 11n^2 + 11n + 3\) then \(Q(n-1) = -Q(-n)\)).

10. What on Earth is Going on Here?

Apéry's incredible proof appears to be a mixture of miracles and mysteries. The dominating question is how to generalise all this, down to the Euler constant \(\gamma\) and up to the general \(\xi(t)\)? Here we have, apparently, the tip of an iceberg which relates \((1 + \sqrt{5})^4\) to \(\xi(3)\) and \(\left(\frac{1}{2}(1+\sqrt{5})\right)^5\) to \(\xi(2)\); we have surprising identities (2) and (2'), and startling continued fractions (produced by Cohen for his Helsinki talk), (7) and (7'). Does the complete berg look like this? For my part I incline to the view that much of what has been presented constitutes a mystification rather than an explanation. For example Richard Askey (Madison, Wisconsin) has pointed out to me that the sequences \{b_n\} and \{b'_n\} may be recognized as special values of certain hypergeometric polynomials; immediately the recurrences...
(2) and (4) become identities relating hypergeometric functions and much of the magic fades away. Unfortunately the difficulties remain, because not all that much is known about the higher generalisations of the classical hypergeometric functions. For this, and other reasons, it is however likely that one should think about recurrences of order greater than 2. This, incidentally, means that the continued fractions constitute a red herring. In any event (7) obscures a fundamental miracle. Its convergents $P_n/Q_n$ are of course such that the sequences $\{P_n\}$ and $\{Q_n\}$ both satisfy

$$U_{n+1} = (34n^3 + 51n^2 + 27n + 5)U_n - n^6U_{n-1}.$$ 

The proof works (not because the continued fraction does not terminate; that only works for regular continued fractions, but) because if $U_0 = 1$, $U_1 = 5$ then it happens that $(n!)^3$ divides the integers $U_n$; more honestly: it is already enough (and is necessary) that for any initial integer values $U_0$, $U_1$, $(n!)^3$ always divides $2[1, \ldots, n]^3U_n$. An analogous miracle makes the recurrence

$$U_{n+1} = (11n^2 + 11n + 3)U_n + n^4U_{n-1}$$ 

useful in proving the irrationality of $\zeta(2)$.\(^9\) These surprises generalise the following quite well known fact (to which I was alerted by Frits Beukers (Leiden)): the recurrence

$$U_{n+1} = (6n + 3)U_n - n^2U_{n-1}$$

is such that $n!$ divides $U_n$ if $U_0 = 1$, $U_1 = 3$; and $n!$ divides $[1, \ldots, n]U_n$ for all integer initial values $U_0$, $U_1$.

Exercise

What are the higher analogues?

8 Show that if

$$B(z) = (1 - 6x + x^2)^{-1/2} = \sum_{n=0}^{\infty} b_n z^n,$$

then the $b_n = \sum_{k=0}^{n} \binom{\frac{n}{k}}{k} (n^k)$, and all are integers. Find an expression for the $a_n$ in

$$A(z) = (1 - 6x + x^2)^{-1/2} \int_0^z (1 - 6t + t^2)^{-1/2} dt = \sum_{n=0}^{\infty} a_n z^n$$

and notice that the $[1, \ldots, n]a_n$ all are integers. Show that sequences $\{a_n\}$ ($a_0 = 0$, $a_1 = 1$) and $\{b_n\}$ ($b_0 = 1$, $b_1 = 3$) both satisfy

$$nu_n + (n - 1)u_{n-2} = (6n - 3)u_{n-1}.$$ 

Now prove that there is a constant $\lambda$ such that

$$A(z) - \lambda B(z) = \sum_{n=0}^{\infty} c_n z^n$$

has no singularity at $3 - 2\sqrt{2}$. Deduce that then $c_n = O(a^{-n})$ with $\alpha = (1 + \sqrt{2})$ and conclude that it follows that log 2 has irrationality degree at most 4, 662 100 831. . .

Of course Exercise (6) should remind us that recurrences may be quite irrelevant to the proof. The vital things then is suitable definition of the $c_n$, etc, one is brought back to looking for generalisations of (2). But, for the present, generalisation of Apéry’s work remains, as they say, a mystery wrapped in an enigma.\(^10\)

Most startling of all though should be the fact that Apéry’s proof has no aspect that would not have been accessible to a mathematician of 200 years ago. The proof we have seen is one that many mathematicians could have found, but missed.

This note was written at Queen’s University, Kingston, Ontario whilst the author was on study leave from the University of New South Wales, Sydney, Australia.

October, 1978

Postscripts. See L. Lewin Dilogarithms and associated functions (Macdonald, London, 1958) for many delightful facts, including the trilogarithm formula of which is given at p. 139. At p. 89 of Louis Comtet Advanced Combinatories (D. Reidel, Dordrecht, 1974) one is astonished to be asked to prove as an exercise that

$$\sum_{n=1}^{\infty} \frac{1}{n^3} = \frac{1}{3} + \frac{2\pi^3}{27} ; \quad \sum_{n=1}^{\infty} \frac{1}{n^{4/2}} = \frac{\pi^3}{9};$$

\(^9\) Tom Cusick (Buffalo) has noticed that the following recurrences also yield continued fractions converging to $\pi^2/6$:

$$n^2u_n = (7n^2 - 7n + 2)u_{n-1} + 8(n - 1)^2u_{n-2}$$

(one solution of which is $\sum_{k=0}^{n} \binom{\frac{n}{k}}{k}$, and

$$n^3u_n = 2(2n - 1) (3n^2 - 3n + 1) u_{n-1} + (4n - 3) (4n - 4)$$

\[(4n - 5) u_{n-2}\]

(A solution is $\sum_{k=0}^{n} \binom{\frac{n}{k}}{k}$).

On first impression the first yields a worse irrationality degree for $\pi^2$ than that obtained by Apéry, and the second does not yield irrationality at all. Apéry’s results are indeed remarkable.

\(^10\) Well, not really. It is just that it is not at all clear where to go. A numerical test (suggested by Cohen) implies that $\zeta(4) = \frac{\pi^4}{90}$ = $(36/17) \sum_{n=1}^{\infty} (1/n^4(2n))$ (so this is true for all practical purposes) and it has been shown by Gosper that

$$\zeta(5) = \frac{5}{2} \sum_{n=1}^{\infty} \left( \frac{1}{2^2} + \frac{1}{2^2} + \cdots + \frac{1}{(n-1)^2} \right) \frac{(-1)^n}{5n^2}.$$ 

David Hawkins (Boulder) suggests similar formulas. Apparently such expressions can be generated virtually at will on using appropriate series accelerator identities.
\[ \sum_{n=1}^{\infty} \frac{1}{n^2 \binom{2n}{n}} = \frac{\pi^2}{18} \; ; \; \sum_{n=1}^{\infty} \frac{1}{n^4 \binom{2n}{n}} = \frac{17\pi^4}{3,240}. \]

Seeing that
\[ \sum_{n=1}^{\infty} \frac{x^{2n}}{n^2 \binom{2n}{n}} = 2 \left( \sin^{-1} \frac{x}{2} \right)^2 \]

(see for example Melzak op cit p. 108) the first three formulæ (and the one with the trilogarithm) become quite accessible to proof, but I had not detected anyone able to prove the expression for \( \xi(4) \), until I proved it in March 1979 after noticing a remark of Lewin that also
\[ \pi^3 \int_0^1 x \left( \log \left( \frac{2 \sin \frac{x}{2} \right) \right)^2 dx = \frac{17\pi^4}{3,240}. \]

Sam Wagstaff (Illinois) and Andrew Odlyzko (Bell Labs) have mentioned to me that numerical evidence suggests that there are formulæ of the shape \( 2 \) or \( 2' \) for \( \xi(t) \) only for \( t = 2, 3, 4 \) and this is verified by my studies in a current manuscript Some wonderful formulæ . . . The recurrences\(^9\) are long known, see Comtet op cit p. 90. One can recognise the \( b_n \) as \( b_n = F_2 \left( \frac{n+1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2} \right) \) and determine the recurrence \( 3 \) by way of three term relations with contiguous balanced series; see J. A. Wilson Hypergeometric series, recurrence relations and some new orthogonal functions (Ph. D. thesis; U. Wisconsin-Madison, 1978).

Frits Beukers (Leiden) A note on the irrationality of \( \xi(2) \) and \( \xi(3) \) (J. Lond. Math. Soc. to appear) has found an elegant approach to Apéry’s proofs which entirely avoids explicit identities, recurrences and other magic. Instead just consider
\[ I = -1/2 \int \int P_n(x)P_n(y) \log xy \frac{dx dy}{1 - xy} = b_n \xi(3) - a_n \]

noticing that the \( b_n \) are integers and the \( a_n \) are rationals with the \( 2 \left[ 1, \ldots, n \right]^3 \) \( a_n \) integral, whilst
\[ |I| \leq \xi(3) \left( 1 - \sqrt{2} \right)^4 n \]

here \( P_n(z) = \frac{1}{n!} \frac{d^n}{dz^n} \left( z^n (1 - z)^n \right) \) is the Legendre polynomial. Again, there is no obvious way to generalise the proof.

In retrospect it seems clear that exercise \( 8 \) really is useful; implications are being considered by Bombieri et al (at Princeton). For example, one’s intuition is just wrong in feeling incredulity at the facts of \( 3 \). All that these report is that the differential equation
\[ \frac{d}{dX} \left( X^4 - 34X^3 + X^2 \right) \frac{d^3 y}{dX^3} + \left( 6X^3 - 103X^2 + 3X \right) \frac{d^2 y}{dX^2} + \left( 7X^2 - 112X + 1 \right) \frac{dy}{dX} + \left( X - 5 \right) y - \left( u_1 - 5u_0 \right) = 0 \]

has two \( G \)-function solutions, namely \( a(X) = 6X + a_2X^2 + \ldots \); \( b(X) = 1 + b_1 X + b_2 X^2 + \ldots \); and \( a(X) - \xi(3) b(X) \) is regular (in fact vanishes) at \( a' = (1 - \sqrt{2})^4 a \). This is interesting, but no longer incredible; and it is readily generalisable . . . All this too is an idea of Beukers. In keeping with the bizarre nature of the events reported here La Recherche No. 97 (France’s Scientific American) contained a report Roger Apéry et l’irrationalité by Michel Mendes-France; the report includes a lively description of the lecture at Marseille (politely suppressed here) although Mendes-France was in the U.S.A. at the time.

Some noxious readers have been critical of my casual use of the \( O \)-symbol; the fault is mine, not Apéry’s. No harm is done. Similarly it has been claimed that Apéry’s proof was not missed by Euler — Euler did not know the prime number theorem’; to me it seems hypercritical to suggest that \( 1, \ldots, n \) \( O((1 + \sqrt{2})^4 n^{3/2}) \) could not have been noticed at the time, had it been needed. Anyhow, I considered it a racy title. It arose after Cohen’s report at Helsinki, with someone sourly commenting ‘A victory for the French peasant . . . ’; to this Nick Katz retorted: ‘No . . ! No! This is marvellous! It is something Euler could have done . . . ’

March 1979

School of Mathematics and Physics
Macquarie University
North Ryde, New South Wales
Australia 2113