Irrationality of ζ(2) and ζ(3): Part 1

Introduction

In 1978, R. Apery gave a mathematical talk which stunned the audience (consisting of other fellow mathematicians). Apery presented a very short proof of the irrationality of $\zeta(3)$ which created utter confusion and many believed his proof to be wrong. However some months later a few other mathematicians (primarily Henri Cohen) verified Apery's proof and concluded that it was correct.

Shortly after all this drama regarding Apery's proof, F. Beukers published another proof of irrationality of $\zeta(3)$ which is much simpler and comprehensible compared to the proof given by Apery. In this series of posts we will provide an exposition of Beukers' Proof. The content of this series is based on Beukers' paper *"A Note on the Irrationality of $\zeta(2)$ and $\zeta(3)$."*

The technique used in the proof applies also to $\zeta(2)$ and the case of $\zeta(2)$ serves to illustrate the technique without getting too complicated. Hence we will start with the case of irrationality of $\zeta(2)$.

Irrationality of $\zeta(2)$

To begin with we provide a definition of the ζ function. Let s be a real number such that $s > 1$ then we define the ζ function by

$$\zeta(s) = \frac{1}{1^s} + \frac{1}{2^s} + \frac{1}{3^s} + \cdots = \sum_{n=1}^{\infty} \frac{1}{n^s} \tag{1}$$

It is known that for even values of s, $\zeta(s)$ is a rational multiple of π^s and hence is transcendental. Therefore the irrationality (and transcendentality) of $\zeta(s)$ is settled for the case when s is an even positive integer. In particular case of $\zeta(2)$ we have $\zeta(2) = \pi^2/6$ and there are many well known proofs of irrationality of π^2.

However when s is odd there is no known simple relation between $\zeta(s)$ and π^s and hence the above approach based on nature of π fails. Apery's (and Beukers' too) achievement was therefore considered to be big as he settled the case for first odd value of $s = 3$ and proved that $\zeta(3)$ is irrational.

Preliminary Results

To understand the proof technique of Beukers we first need to establish some preliminary results on some improper integrals.

Let r, s be non-negative integers with $r > s$. Then we have

$$\int_0^1 \int_0^1 \frac{x^r y^r}{1 - xy}\, dx\, dy = \sum_{n=1}^{\infty} \frac{1}{(n + r)^2} \tag{2}$$

$$\int_0^1 \int_0^1 \frac{x^r y^s}{1 - xy}\, dx\, dy = \frac{1}{r - s}\left\{\frac{1}{s + 1} + \frac{1}{s + 2} + \cdots + \frac{1}{r}\right\} \tag{3}$$

In the above formulas we have to interpret the integrals as being improper because the integrand is not defined for $x = y = 1$. However note that integrals are convergent and all the manipulations regarding them can be justified by taking upper limit as $1 - \epsilon$ and letting $\epsilon \to 0+$.

To establish the first result we observe that

$$\begin{aligned}
\int_0^1 \int_0^1 \frac{x^r y^r}{1 - xy}\, dx\, dy &= \int_0^1 \int_0^1 x^r y^r \left(\sum_{n=0}^{\infty} x^n y^n\right) dx\, dy \\
&= \sum_{n=0}^{\infty} \int_0^1 \int_0^1 x^{n+r} y^{n+r}\, dx\, dy \\
&= \sum_{n=0}^{\infty} \int_0^1 x^{n+r}\, dx \int_0^1 y^{n+r}\, dy \\
&= \sum_{n=0}^{\infty} \frac{1}{n + r + 1} \frac{1}{n + r + 1} \\
&= \sum_{n=1}^{\infty} \frac{1}{(n + r)^2}
\end{aligned}$$

The above result shows that

$$\int_0^1 \int_0^1 \frac{1}{1 - xy}\, dx\, dy = \zeta(2)$$

and for integer $r > 0$ we have

$$\int_0^1 \int_0^1 \frac{x^r y^r}{1 - xy}\, dx\, dy = \zeta(2) - \left(\frac{1}{1^2} + \frac{1}{2^2} + \cdots + \frac{1}{r^2}\right)$$

Next we have

$$\begin{aligned}
\int_0^1\int_0^1 \frac{x^r y^s}{1-xy}\,dx\,dy &= \int_0^1\int_0^1 x^r y^s\left(\sum_{n=0}^{\infty} x^n y^n\right)dx\,dy\\
&= \sum_{n=0}^{\infty}\int_0^1\int_0^1 x^{n+r}y^{n+s}\,dx\,dy\\
&= \sum_{n=0}^{\infty}\int_0^1 x^{n+r}\,dx\int_0^1 y^{n+s}\,dy\\
&= \sum_{n=0}^{\infty}\frac{1}{n+r+1}\frac{1}{n+s+1}\\
&= \sum_{n=0}^{\infty}\frac{1}{r-s}\frac{(n+r-1)-(n+s+1)}{(n+r+1)(n+s+1)}\\
&= \sum_{n=0}^{\infty}\frac{1}{r-s}\left(\frac{1}{(n+s+1)}-\frac{1}{(n+r+1)}\right)\\
&= \frac{1}{r-s}\left\{\frac{1}{s+1}+\frac{1}{s+2}+\cdots+\frac{1}{r}\right\}
\end{aligned}$$

As is evident from their proofs, the formulas (2) and (3) hold even when r, s are positive real numbers, but these integrals can be related to ζ function only when r, s are integers.

Let us define d_n to be the LCM of all numbers $1, 2, \ldots, n$ and $d_0 = 1$. Then it is clear from the above results that

$$\begin{aligned}
\int_0^1\int_0^1 \frac{x^r y^r}{1-xy}\,dx\,dy &= \zeta(2) - \frac{a}{d_r^2}\\
\int_0^1\int_0^1 \frac{x^r y^s}{1-xy}\,dx\,dy &= \frac{b}{d_r^2}
\end{aligned}$$

where r, s are non-negative integers with $r > s$ and a, b are integers. It therefore follows that if $P(x), Q(x)$ are polynomials of degree n with integer coefficients then

$$\int_0^1\int_0^1 \frac{P(x)Q(y)}{1-xy}\,dx\,dy = \frac{a\zeta(2)+b}{d_n^2}$$

where a, b are integers.

Strategy of the Proof

We now choose a specific polynomial $P_n(x)$ defined by

$$P_n(x) = \frac{1}{n!}\frac{d^n}{dx^n}\{x^n(1-x)^n\}$$

It is easy to observe that the polynomial $P(x)$ has integer coefficients and is of degree n (the reason for integer coefficients is that the coefficients are of the form $a/n!$ where a turns out to be a product of n consecutive positive integers). Hence we get

$$I_n = \int_0^1\int_0^1 \frac{(1-y)^n P_n(x)}{1-xy}\,dx\,dy = \frac{a_n\zeta(2)+b_n}{d_n^2} \tag{4}$$

where a_n, b_n are integers dependent on n.

Now our plan is to estimate the integral in the above equation and show that it is non-zero but tends to zero as $n \to \infty$. In fact we need to demonstrate that the integral I_n tends to zero much faster than d_n^{-2} so much so that the whole product $d_n^2 I_n$ also tends to zero as $n \to \infty$. In such a situation the expression $a_n\zeta(2) + b_n$ tends to zero without being equal to zero ever. If $\zeta(2)$ were rational say p/q then we would always have $|a_n\zeta(2) + b_n| \geq 1/q$ and hence it would not tend to zero. This contradiction establishes that $\zeta(2)$ is irrational.

We thus need to establish the following results for the integral I_n defined in equation (4):

1. $I_n \neq 0$ for all positive integers n
2. $d_n^2 I_n \to 0$ as $n \to \infty$

Establishing these results will require the use of integration by parts to transform the integral I_n into a convenient form amenable to reasonable estimation. First we note that if $f(x) = x^n(1-x)^n$ then we have $f^{(j)}(0) = f^{(j)}(1) = 0$ for $j = 0, 1, 2, \ldots, n-1$ and therefore

$$\begin{aligned}
I_n &= \int_0^1 \int_0^1 \frac{(1-y)^n P_n(x)}{1-xy}\, dx\, dy\\
&= \frac{1}{n!}\int_0^1 \int_0^1 \frac{(1-y)^n f^{(n)}(x)}{1-xy}\, dx\, dy\\
&= \frac{1}{n!}\int_0^1 \left\{\left[\frac{(1-y)^n}{1-xy} f^{(n-1)}(x)\right]_{x=0}^{x=1} - \int_0^1 \frac{y(1-y)^n}{(1-xy)^2} f^{(n-1)}(x)\, dx\right\} dy\\
&= -\frac{1}{n!}\int_0^1 \int_0^1 \frac{y(1-y)^n f^{(n-1)}(x)}{(1-xy)^2}\, dx\, dy\\
&= \frac{1\cdot 2}{n!}\int_0^1 \int_0^1 \frac{y^2(1-y)^n f^{(n-2)}(x)}{(1-xy)^3}\, dx\, dy\\
&= (-1)^3\frac{3!}{n!}\int_0^1 \int_0^1 \frac{y^3(1-y)^n f^{(n-3)}(x)}{(1-xy)^4}\, dx\, dy\\
&= (-1)^n\frac{n!}{n!}\int_0^1 \int_0^1 \frac{y^n(1-y)^n f(x)}{(1-xy)^{n+1}}\, dx\, dy
\end{aligned}$$

and thus

$$I_n = (-1)^n \int_0^1 \int_0^1 \frac{y^n(1-y)^n x^n(1-x)^n}{(1-xy)^{n+1}}\, dx\, dy \tag{5}$$

Clearly $I_n \neq 0$ because the integrand is positive for all $x, y \in (0, 1)$.

Estimation of I_n

Now we need to estimate the above integral by estimating the maximum value of the function

$$f(x, y) = \frac{x(1-x)y(1-y)}{1-xy}$$

for $0 \leq x < 1, 0 \leq y < 1$.

Clearly this requires us to solve the equations

$$\frac{\partial f}{\partial x} = 0,\ \frac{\partial f}{\partial y} = 0$$

Since $f(x, y) = f(y, x)$ the relation between x and y induced by $\partial f/\partial x = 0$ is same as the relation between y and x induced by $\partial f/\partial y = 0$ and hence for stationary value of $f(x, y)$ we must have $x = y$ and therefore we need to maximize the function
$g(t) = f(t, t) = (t - t^2)^2/(1 - t^2)$

We have

$$\begin{aligned}&g'(t) = 0\\ &\Rightarrow (1 - t^2)2(t - t^2)(1 - 2t) - (t - t^2)^2(-2t) = 0\\ &\Rightarrow (1 - t^2)(1 - 2t) + t(t - t^2) = 0\\ &\Rightarrow (1 + t)(1 - 2t) + t^2 = 0\\ &\Rightarrow t^2 + t - 1 = 0\\ &\Rightarrow t = \frac{\sqrt{5} - 1}{2}\end{aligned}$$

as $t \in (0, 1)$.

And then we can check (with some labor) that for this value of t we have $g''(t) < 0$ so that the maximum value of $f(x, y)$ is attained at $x = y = (\sqrt{5} - 1)/2$. The maximum value thus obtained is seen to be

$$\begin{aligned}g(t) &= \frac{(t - t^2)^2}{1 - t^2}\\ &= \frac{(t + t - 1)^2}{1 + t - 1}\\ &= \frac{4t^2 - 4t + 1}{t}\\ &= 4t - 4 + \frac{1}{t}\\ &= 4t - 4 + t + 1\\ &= 5t - 3 = \frac{5\sqrt{5} - 11}{2} = \left(\frac{\sqrt{5} - 1}{2}\right)^5\end{aligned}$$

Thus from equation (5) we get

$$\begin{aligned}|I_n| &= \int_0^1\int_0^1 \{f(x, y)\}^n \frac{1}{1 - xy}\,dx\,dy\\ &\leq \left(\frac{\sqrt{5} - 1}{2}\right)^{5n}\int_0^1\int_0^1 \frac{1}{1 - xy}\,dx\,dy\\ &= \zeta(2)\left(\frac{\sqrt{5} - 1}{2}\right)^{5n}\end{aligned} \tag{6}$$

Estimation of d_n

Next we have to estimate d_n. This is achieved by observing that d_n is a multiple of all primes $p \leq n$ with these primes p raised to maximum power k such that $p^k \leq n$. Clearly such maximum value of k is the integer part of $(\log n)/(\log p)$. Therefore we have

$$\begin{aligned}d_n &= \prod_{p\leq n} p^{[\log n/\log p]}\\ &\leq \prod_{p\leq n} p^{\log n/\log p} = \prod_{p\leq n} n = n^{\pi(n)}\end{aligned}$$

where $\pi(n)$ is prime counting function representing numbers of primes less than or equal to n.

From prime number theorem we know that

$$\lim_{n\to\infty}\frac{\pi(n)\log n}{n} = 1$$

so that if A is any fixed number with $A > 1$ then we have

$$\frac{\pi(n)\log n}{n} < A$$

for all sufficiently large values of n. Then we get

$$\pi(n)\log n < nA \Rightarrow n^{\pi(n)} < (e^A)^n = K^n$$

where $K = e^A > e$. Thus if $K > e$ then for all sufficiently large values of n we have $n^{\pi(n)} < K^n$ so that $d_n < K^n$

Conclusion

Combining the above relation with equation (6) we note that

$$d_n^2|I_n| < \zeta(2)\left\{K^2\left(\frac{\sqrt{5}-1}{2}\right)^5\right\}^n$$

Choosing $K = 3$ we can see that the expression inside curly brackets is less than 1 and hence the expression on the right of the above equation (and therefore the expression on the left too) tends to 0 as $n\to\infty$. We have thus established all the results needed to obtain a contradiction (which is arrived by assuming $\zeta(2)$ to be rational). The proof of irrationality of $\zeta(2)$ is thus complete. Note that since $\zeta(2) = \pi^2/6$ the above argument also constitutes another proof of irrationality of π^2 (and therefore of π too).

In the next post we shall deal with the slightly more complicated case of irrationality of $\zeta(3)$.

By Paramanand Singh
Wednesday, October 9, 2013

Labels: Irrational Numbers , Mathematical Analysis