Lecture 20: An application of the PNT

In what follows, we'll use the Prime Number Theorem to prove the irrationality of

$$S(s) = \sum_{k=2}^{\infty} \frac{1}{k^s},$$

established by R. Apery in 1978. (Here I follow a simpler approach due to E. Beukers.)

Let $d_n := \text{lcm}\{1, 2, \ldots, n\}$, $s, r \in \mathbb{N}$ with $r \geq s$.

**Lemma 1:** (a) \[
\int_0^1 \int_0^1 \frac{-\log(xy)}{1-xy} x^{-y} y^s \, dx \, dy = \frac{1}{3} \frac{\pi}{\varpi} \tag{\mathbb{II}}
\]

in case $r > s$

(b) \[
\int_0^1 \int_0^1 \frac{-\log(xy)}{1-xy} x^{-y} y^s \, dx \, dy = 2 \left\{ S(s) - \sum_{k=1}^{\infty} \frac{1}{k^s} \right\}
\]

**Proof:** We claim, for $s > 0$ and $r > s$, that

\[
\int_0^1 \int_0^1 \frac{x^{s+r} y^s}{1-xy} \, dx \, dy = \frac{1}{r-s} \left\{ \frac{1}{s} + \frac{1}{s+1} + \cdots + \frac{1}{r+s} \right\}.
\]

Indeed,

\[
\text{LHS}(s) = \sum_{k=0}^{\infty} \int_0^1 \int_0^1 x^{k+s+r} y^{k+s} \, dx \, dy
\]

\[
= \sum_{k=0}^{\infty} \frac{1}{k^{s+r+1} (k+s+1)(k+s+1)} \tag{\ast\ast}
\]
\[\sum_{k=0}^{r-5} \left\{ \frac{1}{k+s+\sigma+1} - \frac{1}{k+r+\sigma+1} \right\} = \text{RHS}(k),\]

where the switching of \( \Sigma \) and \( \int \) is easily justified on the basis of positivity of the \( x^{k+\sigma+1} \) series (absolute convergence argument).

Differentiating both sides of (\( \star \)) w.r.t. \( \sigma \) gives

\[
\int_0^1 \int_0^1 \log(xy) x^{r+\sigma} y^{s+\sigma} \, dx \, dy = -\frac{1}{r-s} \left\{ \frac{1}{(s+1)^2} + \cdots + \frac{1}{(r+\sigma+1)^2} \right\},
\]

and setting \( \sigma = 0 \) we have

\[
\int_0^1 \int_0^1 \log(xy) x^{r} y^{s} \, dx \, dy = -\frac{1}{r-s} \left\{ \frac{1}{(s+1)^2} + \cdots + \frac{1}{(r)^2} \right\}
\]

\[= \frac{\text{integer}}{(r-s) \cdot (\text{lcm}(s+1, \ldots, r))^2},\]

in which the denominator divides \( d_r^3 \). This gives (a).

Taking \( r = s \) in (\( \star \star \)) gives

\[
\int_0^1 \int_0^1 \frac{x^{r+\sigma} y^{r+\sigma}}{1-x-y} \, dx \, dy = \sum_{k=0}^{r-5} \frac{1}{(k+r+\sigma+1)^2},
\]

and differentiating w.r.t. \( \sigma \) then setting \( \sigma = 0 \) gives

\[
\int_0^1 \int_0^1 \log(xy) x^{r} y^{r} \, dx \, dy = \sum_{k=0}^{r-5} \frac{2}{(k+r+1)^3}
\]

\[= -2 \left\{ \sum_{k=1}^{r} \frac{1}{k^3} - \sum_{k=1}^{r} \frac{1}{k^2} \right\}.\]
Remark: This clearly works for $r = 0$, interpreting \( \frac{\phi}{2^3} \) as 0.

Lemma 2: For $n$ sufficiently large, $d_n < 3^n$.

Proof: $d_n = \prod_{p \leq n} \frac{p^{\log_p n}}{\log_p n} < \prod_{p \leq n} p^{\log_p n} = n^\omega(n)$.

\[\omega(n) = \text{the product of prime powers } \leq n,\]
\[\text{The highest power of } p \text{ which is } \leq n, \text{ is } \log_p n.\]

Let $\epsilon > 0$ be such that $e^{1+\epsilon} < 3$. By the PNT,
\[\exists N \text{ s.t. } n \geq N \Rightarrow \pi(n) < \frac{n}{\log n},\]
\[\Rightarrow \pi(n) \log n < (1+\epsilon) n\]
\[\Rightarrow n \pi(n) < (e^{1+\epsilon})^n < 3^n.\]

In the final proof, we'll use (shifted) Legendre polynomials
\[P_n(x) := \frac{1}{n!} (\frac{d}{dx})^n x^n (1-x)^n \in \mathbb{Z}[x].\]
The first few of these are
\[P_0 = 1, \quad P_1 = 1 - 2x, \quad P_2 = 6x^2 - 6x + 1,\]
\[ P_2 = 20x^3 - 30x^2 + 12x - 1, \]
\[ P_4 = 70x^4 - 140x^3 + 90x^2 - 20x + 1, \text{ etc.} \]

**Theorem** \[ 5(3) \neq 0. \]

**Proof:** Consider

\[ \int_0^1 \int_0^1 -\log(xy) \frac{P_n(x)P_n(y)}{1-xy} \, dx \, dy. \]

Lemma 1 \( \Rightarrow \) \( \#) \Rightarrow \) \( \int_0^1 \int_0^1 -\log(xy) \frac{P_n(x)P_n(y)}{1-xy} \, dx \, dy. \]

Now

\[ \int_0^1 \frac{dz}{1-(1-xy)^2} = -\log \left( \frac{1}{1-xy} \right) \bigg|_{x=0}^1 \]

\[ = -\log(xy) \]

\[ = \frac{-\log(xy)}{1-xy}, \]

so \( \#) \Rightarrow \) \( \int_0^1 \int_0^1 \int_0^1 \frac{P_n(x)P_n(y)}{1-(1-xy)^2} \, dx \, dy \, dz. \]

Next we perform integration by parts in the inner integral:

\[ \int_0^1 \frac{1}{n!(d^n)} \frac{x^n x (1-x)^{n-1}}{1-(1-xy)^2} \, dx = \int_0^1 \frac{\frac{1}{n!(d^n)} x^n (1-x)^n}{1-(1-xy)^2} \, dx \]

\[ n = \frac{1}{\sqrt{(1-xy)^2}}, \quad dv = \frac{1}{n!(d^n)} x^n (1-x)^n \, dx, \quad \text{vanishes at } 0,1. \]
Continuing the process, the successive terms span out a 
\( n! (y^2)^n \) 
which leaves us in the end with
\[
\int_0^1 \frac{1}{x^n(1-x)^m (y^2)^n} \frac{dx}{(1-(1-xy)w)^{m+1}}.
\]
Thus
\[
\Phi(x,y) = \int_0^1 \int_0^1 \int_0^1 \frac{(xy)^n (1-w)^n (1-x)^n P_n(y) \left(\frac{-xy}{1-(1-xy)w}\right)}{(1-(1-xy)w)^{m+1}} \, dy \, dw.
\]
\[
\Phi(x,y) = \int_0^1 \int_0^1 \int_0^1 \frac{(1-w)^n (1-x)^n y^n (1-y)^n (y^2)^n}{(1-(1-xy)w)^{m+1}} \, dy \, dw.
\]
Carry out the same \( n \)-fold \( \int \)-by-parts
for \( y \) as done for \( x \) above.
Consider the function
\[
\phi(x,y,w) = \frac{xwy(1-x)(1-y)(1-w)}{1-(1-xy)w}
\]
on \([0,1]^3\).

Recall this produced (besides
\( x^n(1-x)^m \) \( n(y^2)^n \)).
The partial derivatives must all vanish in order for \( \phi \) to have a maximum. In order for them to vanish somewhere other than the boundary of \([-1,1]^3\), the "interesting" terms in \( \partial_x \phi, \partial_y \phi, \partial_z \phi \) must = 0. These are

\[
\begin{align*}
1 - w - 2x - x^2yw + 2xw &= 0, \\
1 - w - 2y - x^2yw + 2yw &= 0, \\
1 - 2w + w^2 - xyw^2 &= 0.
\end{align*}
\]

The last one says \( \frac{(w-1)^2}{w^2} = xy \). In the first of these, make the substitution \(-x^2yw = -x \frac{(w-1)^2}{w}\), and multiply through by \(w\). This leads to

\[x = \frac{w}{1+w} \quad (= y \text{ by a "symmetry" argument using the 2nd partial term}).\]

So the 3rd partial term becomes

\[0 = 1 - 2w + w^2 - \frac{w^2}{(1+w)^2} w^2,\]

\[= w = \frac{1}{\sqrt{2}} \implies x = \frac{1}{\sqrt{2}+1} = \sqrt{2}-1 \quad (= y).\]

Plugging these back into \( \phi \) gives

\[\phi (\sqrt{2}-1, \sqrt{2}-1, \frac{1}{\sqrt{2}}) = \ldots = (\sqrt{2}-1)^y,\]

which is \( i \) the maximum of \( \phi \) on \([-1,1]^3\).
So \((\#) = \iint_{[0,1]} \frac{(\phi(x,y,w))^n}{1-(1-xy)w} \, dx \, dy \, dw \)

\[
\leq (\sqrt{2}-1)^n \int_{[0,1]} \int_{[0,1]} \frac{dx \, dy \, dw}{1-(1-xy)w} \\
= (\sqrt{2}-1)^n \int_0^1 \int_{-\log \log 2}^{\infty} \frac{-\log \log y}{1-xy} \, dy \\
= (\sqrt{2}-1)^n \cdot 2 \tau(\delta). \\
\text{(Lemma 1(b) + Remark)}
\]

Now from the form \((\#\#)\), we know \((\#) \neq 0\). Hence

\[
0 < \left| A_n + B_n \tau(\delta) \right| d_n^{-3} < 2 \tau(\delta) (\sqrt{2}-1)^n \\
\text{(##)}
\]

\[
\Rightarrow 0 < \left| A_n + B_n \tau(\delta) \right| < 2 \tau(\delta) d_n^{\delta} (\sqrt{2}-1)^n \\
< 2 \tau(\delta) \left( \frac{27}{(\sqrt{2}+1)^2} \right)^n. \\
\text{Lemma 2}
\]

Finally suppose \(\tau(\delta) \in \mathbb{Q}\), i.e.

\((\dagger) \quad \exists P, Q \in \mathbb{N} \text{ s.t. } \tau(\delta) = \frac{P}{Q}.\)
Then, multiplying by $Q$, and taking $n$ sufficiently large, we have

$$0 < |QA_n + B_n P| < 25(3)Q \left( \frac{27}{32} \right)^n < 1$$

which, of course, is impossible, as $QA_n + B_n P \in \mathbb{Z}$.

So (1) cannot hold. □