4 The Irrationality of $\zeta(3)$

For $s > 1$, we define $\zeta(s) = \sum_{n=1}^{\infty} 1/n^s$. We give here a proof by Frits Beukers that $\zeta(3)$ is irrational (the result itself being originally due to R. Apery).

Theorem 10. *The number $\zeta(3) = \sum_{n=1}^{\infty} 1/n^3$ is irrational.*

In addition to Lemma 1 of the previous section (and the notation given there), we make use of the following results.

Lemma 2. *Let r and s be nonnegative integers. If $r > s$, then*

$$\int_0^1 \int_0^1 -\frac{\log(xy)}{1-xy} x^r y^s \, dx \, dy$$

is a rational number whose denominator when reduced divides d_r^3. Also,

$$\int_0^1 \int_0^1 -\frac{\log(xy)}{1-xy} x^r y^r \, dx \, dy = 2\left(\zeta(3) - \sum_{k=1}^{r} \frac{1}{k^3}\right).$$

Proof. Integrating by parts, we obtain that for $k \geq 0$

$$\begin{aligned}
\int_0^1 (\log x) x^{r+k} \, dx &= \lim_{\epsilon \to 0} \int_\epsilon^1 (\log x) x^{r+k} \, dx \\
&= \lim_{\epsilon \to 0} \frac{1}{r+k+1} \int_\epsilon^1 (\log x) \, d\left(x^{r+k+1}\right) \\
&= \lim_{\epsilon \to 0} \frac{-1}{r+k+1} \int_\epsilon^1 x^{r+k} \, dx = \frac{-1}{(r+k+1)^2}.
\end{aligned}$$

We deduce that

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

Integrating now with respect to y in a similar fashion, we obtain

$$\int_0^1 \int_0^1 -\frac{\log(xy)}{1-xy} x^r y^s \, dx \, dy = \sum_{k=0}^{\infty} \left(\frac{1}{(r+k+1)(s+k+1)^2} + \frac{1}{(r+k+1)^2(s+k+1)} \right).$$

If $r > s$, the right-hand side above can be written as

$$\sum_{k=0}^{\infty} \frac{1}{r-s} \left(\frac{1}{(s+k+1)^2} - \frac{1}{(r+k+1)^2} \right) = \frac{1}{r-s} \sum_{k=1}^{r-s} \frac{1}{(s+k)^2}.$$

If $r = s$, we obtain instead

$$2\sum_{k=0}^{\infty} \frac{1}{(r+k+1)^3} = 2\biggl(\zeta(3) - \sum_{k=1}^{r} \frac{1}{k^3}\biggr).$$

Note that for $r > s$ one has $\operatorname{lcm}\bigl((r-s)(s+1)^2, (r-s)(s+2)^2, \ldots, (r-s)r^2\bigr)$ is a divisor of d_r^3. The lemma follows. □

Lemma 3. *Let $D = \{(u, v, w) : u, v, w \in (0, 1)\}$. Then the mapping $f(u, v, w) = (x, y, z)$ defined by $x = u$, $y = v$, and $z = \dfrac{1-w}{1-(1-uv)w}$ from D to D is one-to-one and onto. Furthermore,*

$$\frac{\partial(x, y, z)}{\partial(u, v, w)} = \frac{-uv}{\bigl(1-(1-uv)w\bigr)^2}.$$

Proof. We begin by showing that $f(D) \subseteq D$. Consider $(u, v, w) \in D$. Then $0 < (1-uv)w < w < 1$ so that $1 > 1-(1-uv)w > 1-w > 0$. It follows that

$$1 > \frac{1-w}{1-(1-uv)w} > 0.$$

Hence, $f(u, v, w) \in D$.

Next, we show that $f^2(u, v, w) = (u, v, w)$. This follows from

$$\begin{aligned} f^2(u, v, w) &= f\biggl(u, v, \frac{1-w}{1-(1-uv)w}\biggr) \\ &= \biggl(u, v, \frac{1-(1-uv)w-(1-w)}{1-(1-uv)w} \times \frac{1-(1-uv)w}{1-(1-uv)w-(1-uv)(1-w)}\biggr) \\ &= \biggl(u, v, \frac{1-w+uvw-1+w}{1-w+uvw-1+uv+w-uvw}\biggr) = (u, v, w). \end{aligned}$$

We deduce now that if $(u, v, w) \in D$ and $f(u, v, w) = (u', v', w')$, then $(u', v', w') \in D$ and $f(u', v', w') = (u, v, w)$. It follows that f is one-to-one and onto.

Observe that

$$\begin{aligned} \frac{\partial z}{\partial w} &= \frac{-\bigl(1-(1-uv)w\bigr) + (1-w)(1-uv)}{\bigl(1-(1-uv)w\bigr)^2} \\ &= \frac{-1+w-uvw+1-w-uv+uvw}{\bigl(1-(1-uv)w\bigr)^2} = \frac{-uv}{\bigl(1-(1-uv)w\bigr)^2}. \end{aligned}$$

Hence, for some A and B we have

$$\frac{\partial(x, y, z)}{\partial(u, v, w)} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ A & B & \dfrac{-uv}{\bigl(1-(1-uv)w\bigr)^2} \end{vmatrix} = \frac{-uv}{\bigl(1-(1-uv)w\bigr)^2},$$

which concludes the proof. □

Lemma 4. *The function*

$$f(u,v,w)=\frac{u(1-u)v(1-v)w(1-w)}{1-(1-uv)w}$$

is bounded above by $1/27$ *in the region* D.

Proof. We use the inequality

$$1-(1-uv)w=(1-w)+uvw\geq 2\sqrt{1-w}\sqrt{uvw}.$$

We deduce that for $(u,v,w)\in D$ one has

$$f(u,v,w)\leq\frac{1}{2}\sqrt{u}(1-u)\sqrt{v}(1-v)\sqrt{w(1-w)}.$$

For $t\in[0,1]$, the maximum of $t(1-t^2)$ occurs at $t=1/\sqrt{3}$ and the maximum of $t(1-t)$ occurs at $t=1/2$. Hence, we deduce that

$$f(u,v,w)\leq\frac{1}{2}\times\frac{1}{\sqrt{3}}\left(1-\frac{1}{3}\right)\times\frac{1}{\sqrt{3}}\left(1-\frac{1}{3}\right)\times\sqrt{\frac{1}{2}\left(1-\frac{1}{2}\right)}=\frac{1}{27}.$$

This establishes the lemma. □

Proof of Theorem 10. Consider the integral

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

where P_n is as defined in the proof of Theorem 8. Recall $P_n(x)$ is in $\mathbb{Z}[x]$ and of degree n so that $P_n(x)P_n(y)$ is a sum of terms of the form $a_{ij}x^iy^j$ where $0\leq i\leq n$, $0\leq j\leq n$, and $a_{ij}\in\mathbb{Z}$ for all such i and j. Also, $a_{ii}\geq 0$ for each i and $a_{nn}>0$ (in fact, $a_{ii}>0$ for each i, but this is not needed). Lemma 2 implies that the double integral above is a sum of rational numbers whose denominators divide d_n^3 plus a positive integral multiple of $\zeta(3)$. In particular, there exist integers A_n and B_n with $B_n>0$ such that the double integral equals $\big(A_n+B_n\zeta(3)\big)/d_n^3$.

We now find a second expression for the double integral. Since

$$-\frac{\log(xy)}{1-xy}=\int_0^1\frac{1}{1-(1-xy)z}\,dz,$$

the double integral becomes

$$\int_0^1\int_0^1\int_0^1\frac{P_n(x)P_n(y)}{1-(1-xy)z}\,dz\,dx\,dy.$$

Note that for $0\leq k\leq n-1$ the multiple derivative

$$\frac{d^k}{dx^k}\big(x^n(1-x)^n\big)$$

can be expressed as a sum of terms each having both x and $1-x$ as a factor. Switching order of integration and integrating by parts repeatedly, we obtain

$$\int_0^1\int_0^1 -\frac{\log(xy)}{1-xy}P_n(x)P_n(y)\,dx\,dy = \int_0^1\int_0^1\int_0^1 \frac{P_n(x)P_n(y)}{1-(1-xy)z}\,dx\,dy\,dz$$

$$= \frac{1}{n!}\int_0^1\int_0^1\int_0^1 P_n(y)\frac{\frac{d^n}{dx^n}\big(x^n(1-x)^n\big)}{1-(1-xy)z}\,dx\,dy\,dz$$

$$= \frac{1}{n!}\int_0^1\int_0^1\int_0^1 P_n(y)\frac{1}{1-(1-xy)z}\,d\bigg(\frac{d^{n-1}}{dx^{n-1}}\big(x^n(1-x)^n\big)\bigg)\,dy\,dz$$

$$= \frac{1}{n!}\int_0^1\int_0^1\int_0^1 P_n(y)yz\frac{\frac{d^{n-1}}{dx^{n-1}}\big(x^n(1-x)^n\big)}{\big(1-(1-xy)z\big)^2}\,dx\,dy\,dz$$

$$= \cdots = \frac{1}{n!}\int_0^1\int_0^1\int_0^1 P_n(y)n!(yz)^n\frac{x^n(1-x)^n}{\big(1-(1-xy)z\big)^{n+1}}\,dx\,dy\,dz$$

$$= \int_0^1\int_0^1\int_0^1 \frac{x^ny^nz^n(1-x)^nP_n(y)}{\big(1-(1-xy)z\big)^{n+1}}\,dx\,dy\,dz.$$

We apply now the transformation of Lemma 3 so that

$$z^n = \frac{(1-w)^n}{\big(1-(1-uv)w\big)^n}$$

and

$$\big(1-(1-xy)z\big)^{n+1} = \bigg(1-(1-uv)\frac{(1-w)}{1-(1-uv)w}\bigg)^{n+1} = \frac{(uv)^{n+1}}{\big(1-(1-uv)w\big)^{n+1}}.$$

The above integral becomes

$$\int_0^1\int_0^1\int_0^1 \frac{u^nv^n(1-w)^n(1-u)^nP_n(v)\big(1-(1-uv)w\big)^{n+1}}{\big(1-(1-uv)w\big)^n(uv)^{n+1}}\cdot\frac{uv}{\big(1-(1-uv)w\big)^2}\,du\,dv\,dw$$

$$= \int_0^1\int_0^1\int_0^1 (1-w)^n(1-u)^n\frac{P_n(v)}{1-(1-uv)w}\,du\,dv\,dw.$$

By similarity of this integral with a previous integral, exchanging the order of integration and integrating by parts with respect to v and finally changing the order of integration back gives

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

$$= \int_0^1\int_0^1\int_0^1 u^n(1-u)^nv^n(1-v)^nw^n(1-w)^n\frac{du\,dv\,dw}{\big(1-(1-uv)w\big)^{n+1}}.$$

We apply Lemma 4 and then Lemma 2 to obtain

$$0 < \int_0^1 \int_0^1 -\frac{\log(xy)}{1-xy} P_n(x)P_n(y)\,dx\,dy \leq (1/27)^n \int_0^1 \int_0^1 \int_0^1 \frac{du\,dv\,dw}{1-(1-uv)w}$$
$$= (1/27)^n \int_0^1 \int_0^1 -\frac{\log(uv)}{1-uv}\,du\,dv = 2(1/27)^n \zeta(3).$$

We now deduce that

$$0 < \frac{|A_n + B_n\zeta(3)|}{d_n^3} < 2\zeta(3)(1/27)^n$$

for an arbitrary positive integer n and some integers A_n and B_n. Assume now that $\zeta(3) = a/b$ for some integers a and b with $b > 0$. By Lemma 1,

$$0 < |bA_n + aB_n| \leq 2\zeta(3)(1/27)^n d_n^3 b$$
$$< 2\zeta(3)(1/27)^n 2.8^{3n} b = 2\zeta(3)\big(2.8^3/27\big)^n b < 2\zeta(3)(0.9)^n b.$$

Since $bA_n + aB_n$ is an integer, we obtain a contradiction for n sufficiently large (so that $(0.9)^n < 1/(2\zeta(3)b)$). Hence, $\zeta(3)$ is irrational. □

Homework:

1. Show that the argument above can be modified (actually simplified) to establish the irrationality of $\zeta(2) = \sum_{n=1}^{\infty} 1/n^2$. Use the identity

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

   Besides giving some indication as to how the above identity can be established, you should provide lemmas similar to Lemma 2 and Lemma 4 (but may need to give a little thought as to what these should be).

2. Let $f(u,v,w)$ be as in Lemma 4. Prove that the maximum value of $f(u,v,w)$ in the region D is $(\sqrt{2}-1)^4$.
