A2

Orthogonality and Irrationality

Overview

This appendix is an application of orthogonalization of particular Müntz systems to the proof of the irrationality of $\zeta(3)$ and some other familiar numbers. It reproduces Apéry's remarkable proof of the irrationality of $\zeta(3)$ in the context of orthogonal systems.

Orthogonality and Irrationality

Apéry's wonderful proof of the irrationality of $\zeta(3)$ amounts to showing that

$$0 < |d_n^3 a_n \zeta(3) - b_n| \to 0\,,$$

where b_n is an integer,

$$a_n := \sum_{k=0}^{n} \binom{n+k}{k}^2 \binom{n}{k}^2\,,$$

and

$$d_n := \operatorname{lcm}\{1, 2, \ldots, n\}\,.$$

Here lcm denotes the least common multiple; see van der Poorten [79] and Beukers [79]. In [81] Beukers recast the proof using Padé approximations.

Many, maybe most, irrationality proofs may be based on approximation by Padé approximants and related orthogonal polynomials; see, for example, Borwein [91a], [92] or Chudnovsky and Chudnovsky [84]. It is the intention of this appendix to try to put the proof of the irrationality of $\zeta(3)$ into the framework of orthogonality. From the general orthogonalization of the system

$$(x^{\lambda_0}, x^{\lambda_1}, x^{\lambda_2}, \dots)$$

on $[0,1]$, where the numbers λ_j are nonnegative and distinct, specializing to the case when

$$\lambda_{2j} = j \qquad \text{and} \qquad \lambda_{2j+1} = j + \epsilon, \qquad j = 0, 1, \dots$$

where ϵ decreases to 0, is a very natural thing to do. This is how one should interpret orthogonalizing the system $(x^0, x^0, x^1, x^1, \dots)$. This leads to orthogonal functions that generalize the Legendre polynomials and are of the form

$$p_n(x) \log x + q_n(x)$$

with polynomials $p_n, q_n \in \mathcal{P}_n$ of degree n. Legendre polynomials are closely tied to irrationality questions concerning log (see Borwein and Borwein [87], Chapter 11), and higher-order analogs prove to be the basis of dealing with the irrationality of the trilog $\left(\sum_{n=1}^{\infty} x^n/n^3\right)$ for some values of x. We think that the proof of the irrationality of $\zeta(3)$ flows quite naturally from this point of view. Although in the end (Lemma A.2.3) we get back to Beukers' integral approach to the irrationality of $\zeta(3)$ (as indeed we must). What follows, up to one application of the prime number theorem, is a self-contained proof of the irrationality of $\zeta(3)$.

The orthogonalization in question is the content of the first theorem.

Theorem A.2.1. *Let*

$$G_n(x) := \frac{1}{2\pi i} \int_{\Gamma} \frac{\prod_{k=0}^{n-1} (t+k+1)^2}{\prod_{k=0}^{n} (t-k)^2} (t+n+1) x^t \, dt \,,$$

where Γ is any simple contour containing $t = 0, 1, \dots, n$. Then

$$G_n(x) = p_n(x) \log x + q_n(x) \,,$$

where

$$p_n(x) = \sum_{k=0}^{n} \binom{n}{k}^2 \binom{n+k}{k}^2 (n+k+1)\, x^k \,.$$

Furthermore, we have the orthogonality relations

$$\int_0^1 G_n(x) x^k \, dx = 0 \,, \qquad k = 0, 1, \dots, n \,;$$

$$\int_0^1 G_n(x) (\log x) x^k \, dx = 0 \,, \qquad k = 0, 1, \dots, n-1 \,;$$

and

$$\int_0^1 G_n^2(x)\,dx = \frac{1}{2n+1}\,.$$

Proof. As in Section 3.4, the representation of G_n is just the evaluation of the integral at the poles, $t = 0, 1, \ldots, n$, by the residue theorem. The proof of the orthogonality conditions is a straightforward exercise on evaluating

$$\int_0^1 x^k (\log x)^j G_n(x)\,dx$$

by interchanging the order of integration as in the proof of Theorem 3.4.3. □

We need to modify these forms marginally to give a zero at 1, as in the next result.

Theorem A.2.2. *Let G_n and Γ be defined as in Theorem A.2.1. Then*

$$\begin{aligned} F_n(x) &:= \frac{1}{x^{n+1}} \int_0^x y^n G_n(y)\,dy \\ &= \frac{1}{2\pi i} \int_\Gamma \frac{\prod_{k=0}^{n-1}(t+k+1)^2}{\prod_{k=0}^{n}(t-k)^2} x^t\,dt = A_n(x)\log x + B_n(x)\,, \end{aligned}$$

where

$$A_n(x) := \sum_{k=0}^{n} \binom{n+k}{k}^2 \binom{n}{k}^2 x^k$$

and

$$B_n(x) := \sum_{k=0}^{n} c_k x^k$$

with

$$c_k := \binom{n+k}{k}^2 \binom{n}{k}^2 \left\{ \sum_{i=0}^{n-1} \frac{2}{k+i+1} - \sum_{\substack{i=0 \\ i\neq k}}^{n} \frac{2}{k-i} \right\}.$$

Furthermore,

$$\int_0^1 F_n(x) F_m(x)\,dx = \frac{2\delta_{n,m}}{(2n+1)^3}\,, \qquad \delta_{n,m} := \begin{cases} 0 & \text{if } n \neq m \\ 1 & \text{if } n = m\,; \end{cases}$$

$$\int_0^1 F_n(x) x^k (\log x)^j\,dx = 0\,, \qquad k = 0, 1, \ldots, n-1\,, \quad j = 0, 1\,;$$

and

$$F_n(1) = 0\,, \qquad F_n'(1) = 1\,, \qquad F_n''(1) = 2n^2 + 2n - 1\,.$$

Proof. This follows much as in Theorem A.2.1. The fact that $F_n(1) = 0$ is just the statement that G_n is orthogonal to x^n. □

The fact that $F_n(1) = 0$ is critical in what follows. It immediately gives that $B_n(x)$ has a zero at 1, so $B_n(x)/(1-x)$ is a polynomial, as in part c] of the following corollary. (This is the only part of the corollary we need, but the corollary is of some interest in its own right.)

Corollary A.2.3. *Let F_n be defined as in Theorem A.2.2. Then*

a] *F_n has $2n+1$ zeros on $(0,1]$.*

b] *$A_n(x)$ has all real negative zeros.*

c] *$B_n(x)/(1-x)$ is a polynomial with all real negative zeros that interlace the zeros of $A_n(x)$.*

Proof. The orthogonality conditions give $2n$ zeros of F_n on $(0,1)$ in a standard fashion, and there is one zero at 1. The real negative zeros of $A_n(x)$ and $B_n(x)$ and their interlacing follow from the fact that

$$\log x + \frac{B_n(x)}{A_n(x)}$$

has $2n+1$ zeros on $(0,1]$, and known results on interpolating Stieltjes transforms by rational functions (see, for example, Baker and Graves Morris [81] or Borwein [83]). □

One can, from the integral representation, deduce the next corollary, which is also not actually needed in the proof of the irrationality of $\zeta(3)$.

Corollary A.2.4. *Both F_n and A_n satisfy*

$$x^2(y_n'' - y_{n-1}'') - 2nx(y_n' + y_{n-1}') + x(y_n' - y_{n-1}') + n^2(y_n - y_{n-1}) = 0\,.$$

From Theorem A.2.2 we obtain (see E.2 b]) the following:

Corollary A.2.5. *Let $d_n := \operatorname{lcm}\{1, 2, \ldots, n\}$. Then the polynomial $d_n B_n(x)$ has integer coefficients.*

We get an approximation to $\zeta(3)$ by integration over the unit square.

Theorem A.2.6.

$$-\frac{1}{2}\int_0^1 \int_0^1 \frac{F_n(xy)}{1-xy}\,dx\,dy = A_n(1)\zeta(3) + R_n\,,$$

where $2d_n^3 R_n$ is an integer.

Proof. Recall that

$$-\frac{1}{2}\int_0^1\int_0^1 \frac{\log(xy)}{1-xy} = \zeta(3)$$

since

$$\int_0^1\int_0^1 (xy)^n \log(xy)\,dx\,dy = \frac{-2}{(n+1)^3}\,.$$

We have

$$\frac{F_n(xy)}{1-xy} = \frac{A_n(xy)-A_n(1)}{1-xy}\log(xy) + \frac{A_n(1)}{1-xy}\log(xy) + \frac{B_n(xy)}{1-xy}\,.$$

The first term of the last expression is a polynomial multiple of $\log(xy)$, and the last term is a polynomial. Both have degree $n-1$ in xy. Here we use part c] of Corollary A.2.3 in an essential way. Integrating the above equation with respect to x on $[0,1]$ and with respect to y on $[0,1]$, we get the identity of the theorem. The fact that $d_n^3 R_n$ is an integer can be seen as follows. One d_n arises from each of the two integrations of a polynomial of degree $n-1$ with integer coefficients and one d_n comes from Corollary A.2.5. □

Theorem A.2.7. *The number $\zeta(3)$ is irrational.*

Proof. It now suffices to show that there is an $\epsilon > 0$ for which

$$0 < \left|\int_0^1\int_0^1 \frac{F_n(xy)}{1-xy}\,dx\,dy\right| = O\left(\frac{1}{e^{(3+\epsilon)n}}\right) = o\left(\frac{1}{d_n^3}\right)$$

since by the prime number theorem, $\operatorname{lcm}\{1,2,\ldots,n\} = O(e^{(1+\epsilon)n})$ for every $\epsilon > 0$; see Borwein and Borwein [87, p. 377]. (We use the notation $b_n = o(a_n)$ if $b_n = \epsilon_n a_n$ with $\lim_{n\to\infty}\epsilon_n = 0$.) This can be proven in a number of ways, we chose to connect this proof via Padé approximants to the integral estimate due to Beukers. This is the content of the following results and, in particular, Lemma A.2.10. From Lemma A.2.10, the above estimates are easy since the integrand in the right-hand side in Lemma A.2.10 satisfies

$$0 < \frac{xyv(1-x)(1-y)(1-v)}{1-(1-xy)v} \le (\sqrt{2}-1)^4$$

on the open unit cube $0 < x, y, v < 1$. □

The following lemma gives standard representations for the Padé approximants to log and can be checked by expanding the integrals; see E.3.

Lemma A.2.8 (Padé Approximants). *We have*

$$\frac{(n!)^2}{2\pi i}\int_\Gamma \frac{x^t\,dt}{\prod_{k=0}^n (t-k)^2} = (x-1)^{2n+1}\int_0^1 \frac{v^n(1-v)^n\,dv}{(1-(1-x)v)^{n+1}}$$
$$= p_n(x)\log x + q_n(x) = O((x-1)^{2n+1})\,,$$

where Γ is a simple contour containing $t = 0, 1, \ldots, n$, and p_n and q_n are polynomials of degree n. (So p_n/q_n is the (n,n) Padé approximant to $\log x$ at 1.)

Lemma A.2.9 (Rodrigues-Type Formula). *With Γ as in Theorem A.2.1,*

$$F_n(xy) = \frac{d^n}{dy^n}\frac{d^n}{dx^n}\frac{x^n y^n}{2\pi i}\int_\Gamma \frac{(xy)^t\,dt}{\prod_{k=0}^n (t-k)^2}$$
$$= \frac{1}{(n!)^2}\frac{d^n}{dy^n}\frac{d^n}{dx^n}\int_0^1 \frac{(xy)^n v^n(1-v)^n(xy-1)^{2n+1}}{(1-(1-xy)v)^{n+1}}\,dv\,.$$

Proof. This follows from differentiating the two representations derived from Lemma A.2.8 coupled with Theorem A.2.2. □

Lemma A.2.10. *We have*

$$\int_0^1\int_0^1 \frac{F_n(xy)}{1-xy}\,dx\,dy$$
$$= -\int_0^1\int_0^1\int_0^1 \frac{[xyv(1-x)(1-y)(1-v)]^n\,dx\,dy\,dv}{(1-(1-xy)v)^{n+1}}\,.$$

Proof. For k, n nonnegative integers

$$\int_0^1\int_0^1 (xy)^{n+k}(1-x)^n(1-y)^n\,dx\,dy$$
$$= -\frac{1}{(n!)^2}\int_0^1\int_0^1 \frac{1}{1-xy}\frac{d^n}{dy^n}\frac{d^n}{dx^n}(xy)^{n+k}(xy-1)^{2n+1}\,dx\,dy\,.$$

(Both sides equal

$$\left(\frac{n!(n+k)!}{(2n+k+1)!}\right)^2\,,$$

though this is not completely transparent; see E.4.) So

$$\frac{1}{(n!)^2}\int_0^1\int_0^1 \frac{1}{1-xy}\frac{d^n}{dy^n}\frac{d^n}{dx^n}x^n y^n(xy-1)^{2n+1+k}\,dx\,dy$$
$$= -\int_0^1\int_0^1 [xy(1-x)(1-y)]^n\,(xy-1)^k\,dx\,dy\,.$$

Hence, on expanding $(1-(1-xy)v)^{-n-1}$ in the following integrands,

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

and with Lemma A.2.9 we are done. □

Comments, Exercises, and Examples.

The approach of this appendix follows Borwein, Dykshoorn, Erdélyi, and Zhang [to appear]. Beukers' [79] very elegant recasting of Apéry's proofs of the irrationality of $\zeta(2)$ and $\zeta(3)$ is also presented in Borwein and Borwein [87]. E.5 recasts the irrationality of $\zeta(2) = \frac{1}{6}\pi^2$ into a form similar to the proof of the irrationality of $\zeta(3)$. E.6 treats the irrationality of $\log 2$. Mahler [31] casts transcendence results for exp and log in terms of general systems of Padé approximants.

E.1 Proof of Corollaries A.2.4 and A.2.5.

a] Prove Corollary A.2.4 from the explicit representations of Theorem A.2.2.

b] Prove Corollary A.2.5 from the formula for c_k in Theorem A.2.2.

Hint: Observe that if p is a prime and $n < p^\alpha \le n+k$ for some integer α, then p divides $\binom{n+k}{k}$. This is fairly straightforward from Euler's formula for the largest power of a prime dividing a factorial. □

E.2 Formulas for $\zeta(n)$.

a] Show that

$$\int_0^1\cdots\int_0^1\int_0^1 \frac{dx_1\,dx_2\cdots dx_n}{1-x_1x_2\cdots x_n} = \zeta(n) := \sum_{k=1}^{\infty}\frac{1}{k^n}\,.$$

b] Show that

$$\int_0^1 \frac{\log x}{1-x}\,dx = -\frac{\pi^2}{6}\,.$$

c] Find a closed formula for

$$\int_0^1\cdots\int_0^1\int_0^1 \frac{\log(x_1x_2\cdots x_n)}{1-x_1x_2\cdots x_n}\,dx_1\,dx_2\cdots dx_n$$

in terms of $\zeta(n+1)$.

E.3 Proof of Lemma A.2.8. Give a proof of Lemma A.2.8.

E.4 The Identity in the Proof of Lemma A.2.10. Prove the identity

$$\begin{aligned}\int_0^1 \int_0^1 (xy)^{n+k}(1-x)^n(1-y)^n\,dx\,dy \\ &= -\frac{1}{(n!)^2}\int_0^1\int_0^1 \frac{1}{1-xy}\frac{d^n}{dy^n}\frac{d^n}{dx^n}(xy)^{n+k}(xy-1)^{2n+1}\,dx\,dy \\ &= \left(\frac{n!(n+k)!}{(2n+k+1)!}\right)^2,\end{aligned}$$

where k and n are nonnegative integers. Follow parts a] to c].

Let k and n be nonnegative integers.

a] Show that

$$A_{n,k} := \int_0^1\int_0^1 (xy)^{n+k}(1-x)^n(1-y)^n\,dx\,dy = \left(\frac{n!(n+k)!}{(2n+k+1)!}\right)^2.$$

b] Let

$$B_{n,k} := -\frac{1}{(n!)^2}\int_0^1\int_0^1 \frac{1}{(1-xy)}\frac{d^n}{dy^n}\frac{d^n}{dx^n}x^{n+k}(xy-1)^{2n+1}\,dx\,dy\,.$$

Show that

$$\begin{aligned}n^2 B_{n,k} =&[(n+k)^2+(2n+1)(2n+2k+1)+(2n+1)(2n)]B_{n-1,k+2} \\ &-[2(n+k)^2+(2n+1)(2n+2k+1)]B_{n-1,k+1} \\ &+(n+k)^2 B_{n-1,k}\,.\end{aligned}$$

c] Show that $A_{0,k} = B_{0,k}$. Show that the values

$$A_{n,k} = \left(\frac{n!(n+k)!}{(2n+k+1)!}\right)^2$$

satisfy the recurrence relation established by part b] for the values $B_{n,k}$.

E.5 The Irrationality of π^2. Consider

$$F_n(x) := \frac{n!}{(2\pi i)}\int_\Gamma \frac{\prod_{k=1}^n (t+k)}{\prod_{k=0}^n (t-k)^2}\,x^t\,dt\,,$$

where Γ is any simple contour containing the poles at $0, 1, \ldots, n$.

a] Show that

$$\int_0^1 \frac{F_n(x)}{1-x}\,dx = \sum_{k=0}^{n} \binom{n+k}{k} \binom{n}{k}^2 \zeta(2) + R_n\,,$$

where $d_n^2 R_n$ is an integer ($d_n := \operatorname{lcm}\{1, 2, \ldots, n\}$ as before).

Hint: Show that

$$F_n(x) = A_n(x)\log x + B_n(x)\,,$$

where

$$A_n(x) := \sum_{k=0}^{n} \binom{n+k}{k}^2 \binom{n}{k}^2 x^k$$

and

$$B_n(x) := \sum_{k=0}^{n} c_k x^k$$

with

$$c_k := \binom{n+k}{k} \binom{n}{k}^2 \left\{ \sum_{i=1}^{n} \frac{1}{k+i} - \sum_{\substack{i=0 \\ i\neq k}}^{n} \frac{2}{k-i} \right\}.$$

Write

$$\frac{F_n(x)}{1-x} = \frac{A_n(x) - A_n(1)}{1-x}\,\log x + \frac{A_n(1)}{1-x} + \frac{B_n(x)}{1-x}\,.$$

Show that $F_n(1) = 0$, hence $B_n(1) = 0$. Recall that

$$-\int_0^1 \frac{\log x}{1-x\,dx} = \zeta(2)\,.$$

Now the conclusions follow, as in the proof of Theorem A.2.6. □

b] Show that

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

Hint: Use Lemma A.2.8. □

c] Show that there exists a constant c independent of n such that

$$0 < \left| \int_0^1 \frac{F_n(x)}{1-x}\,dx \right| \le c \left(\frac{\sqrt{5}-1}{2} \right)^{5n}$$

and deduce that $\zeta(2) = \frac{1}{6}\pi^2$ irrational. Hence π is irrational.

E.6 The Irrationality of $\log 2$. Let

$$p_n(x) := \frac{1}{n!}\frac{d^n}{dx^n}[x^n(1-x)^n] = \sum_{k=0}^{n} \binom{n+k}{k}\binom{n}{k}(-1)^k x^k$$

be the nth Legendre polynomial on $[0,1]$.

a] Show that

$$\begin{aligned}\int_0^1 \frac{p_n(x)}{1+x}\,dx &= \int_0^1 \frac{p_n(x)-p_n(-1)}{1+x}\,dx + \int_0^1 \frac{p_n(-1)}{1+x}\,dx \\ &= p_n(-1)\log 2 + R_n\,,\end{aligned}$$

where $d_n R_n$ is an integer ($d_n := \operatorname{lcm}\{1,2,\ldots,n\}$ as before).

b] Show that

$$\left|\int_0^1 \frac{p_n(x)}{1+x}\,dx\right| = \int_0^1 \frac{x^n(1-x)^n}{(1+x)^{n+1}}\,dx\,.$$

c] Use parts a] and b] to show that $\log 2$ is irrational.