Irrationality proofs using modular forms

Re-TEXed by Seewoo Lee*

Frits Beukers

Last updated: November 26, 2024

0 Introduction

In the years following Apéry's discovery of his irrationality proofs for $\zeta(2)$, $\zeta(3)$ (see [8]), it has become clear that these proofs do not only have significance as irrationality proofs, but the numbers that occur in them serve as interesting examples for several phenomena in algebraic geometry and modular form theory. See [1, 2, 3, 5] for congruences of the Apéry numbers and [4, 7] for geometrical and modular interpretations.[1] Furthermore, it turns out that Apery's proofs themselves are in fact simple consequences of elementary complex analysis on spaces of certain modular forms. In the present paper we describe this analysis together with some generalisations in Theorems 1 to 5. For example, we prove that $8\zeta(3) - 5\sqrt{5}L(3) \notin \mathbb{Q}(\sqrt{5})$, where $L(3) = \sum_{n=1}^{\infty} \left(\frac{n}{5}\right) n^{-3}$. Although the use of modular forms in irrationality proofs looks promising at first sight, the yield of new irrationality results thus far is disappointingly low. However, in methods such as this it is easy to overlook some simple tricks that may give new interesting results.

The first section of this paper describes the general framework of the proofs. This section may seem vague at first sight, but in combination with the proof of Theorem 1 we hope that things will be clear. We have given the proof of Theorem 1 as extensively as possible in order to set it as an example for the other proofs, where we omit some minor details now and then.

*seewoo5@berkeley.edu. Some typos are fixed with footnotes.

[1] The original citations were in a different order, but it seems that this is the correct order.

1 Preliminaries

In this section we shall describe the general principles which are used in the arguments of the following sections.

Let $t(q) = \sum_{n=0}^{\infty} t_n q^n$ a power series convergent for all $|q| < 1$. Let $w(q)$ be another analytic function on $|q| < 1$. We like to study w as function of t. In general it will be a multivalued function over which we have no control. However, we shall introduce some assumptions. First, $t_0 = 0, t_1 \neq 0$. Let now $q(t)$ be the local inverse of $t(q)$ with $q(0) = 0$. Choose $w(q(t))$ for the value of w around $t = 0$. In order to determine the radius of convergence of the powerseries $w(q(t)) = \sum_{n=0}^{\infty} w_n t^n$ we introduce branching values of t. We say that t branches above t_0, if either t_0 is not in the image of t, or if $t'(q_0) = 0$ for some q_0 with $t(q_0) = t_0$. In other words, t branches above t_0, if the map $t : \{|q| < 1\} \to \mathbb{C}$ is not a local covering above t_0. We call such a t_0 a branching value of t. Now assume, that t has a discrete set of branching values $t_1, t_2, \ldots$ where we have excluded zero as a possible value and suppose $|t_1| < |t_2| < \ldots$. It is clear now that the radius of convergence is in general $|t_1|$. We shall be interested in cases where the radius of convergence is larger than $|t_1|$. Let γ be a closed contour in the complex t-plane beginning and ending at the origin, not passing through any t_i and which encircles the point t_1 exactly once. Suppose that analytic continuation of $w(q(t))$ along γ again yields the same branch of $w(q(t))$. Then $w(q(t))$ can be continued analytically to the disc $|t| < |t_2|$ with exception of the possible isolated singularity t_1. If $w(q(t))$ remains bounded around we can conclude that the radius of convergence is at least t_2. Our irrationality proofs consist exactly of the construction of such instances. The point of having a radius of convergence as large as possible consists of the following Proposition.

Proposition 1.1. Let $f_0(t), f_1(t), \ldots, f_k(t)$ be power series in t. Suppose that for any $n \in \mathbb{N}$, $i = 0, 1, \ldots, k$ the n-th coefficient in the Taylor series of is rational and has denominator dividing $d^n[1, \ldots, n]^r$ where r, d are certain fixed positive integers and $[1, \ldots, n]$ is the lowest common multiple of $1, \ldots, n$. Suppose there exist real numbers $\theta_1, \ldots, \theta_k$ such that $f_0(t) + \theta_1 f_1(t) + \cdots + \theta_k f_k(t)$ has radius of convergence ρ and infinitely many nonzero Taylor coefficients. If $\rho > de^r$, then at least one of $\theta_1, \ldots, \theta_k$ is irrational.

Remark. Note that if $k = 1$ we have an honest irrationality proof.

Proof. Choose $\epsilon > 0$ such that $\rho - \epsilon > de^{r(1+\epsilon)}$. Let $f_i = \sum_{n=0}^{\infty} a_{i,n} t^n$. Since the radius of the convergence of $f_0 + \theta_1 f_1 + \cdots + \theta_k f_k$ is ρ, we have for sufficiently large n, $|a_{0,n}+\theta_1 a_{1,n}+\cdots+\theta_k a_{k,n}| \le (\rho-\epsilon)^{-n}$. Suppose $\theta_1,\ldots,\theta_k$ are all rational and have common denominator D. Then $A_n = Dd^n[1,\ldots,n]^r|a_{0,n}+\theta_1 a_{1,n}+\cdots+\theta_k a_{k,n}|$ is an integer smaller than $Dd^n[1,\ldots,n]^r(\rho-\epsilon)^{-n}$. By the prime number theorem we have $[1,\ldots,n] < e^{(1+\epsilon)n}$ for sufficiently large n, hence $A_n < D(de^{r(1+\epsilon)}(\rho-\epsilon)^{-1})^n$. Since $de^{r(1+\epsilon)}(\rho-\epsilon)^{-1} < 1$ this implies that $A_n = 0$ for sufficiently large n, in contradiction with the assumption $A_n \neq 0$ for infinitely many n. Thus our proposition follows. □

The construction of the functions $t(q)$ and $w(q)$ will proceed using modular forms and functions. The values for which we obtain irrationality results are in fact values at integral points of Dirichlet series associated to modular forms.

Proposition 1.2. Let $F(\tau) = \sum_{n=1}^{\infty} a_n q^n$, $q = e^{2\pi i\tau}$ be a Fourier series convergent for $|q| < 1$, such that for some $k, n \in \mathbb{N}$,

$$F\left(-\frac{1}{N\tau}\right) = \varepsilon(-i\tau\sqrt{N})^k F(\tau)$$

where $\varepsilon = \pm 1$. Let $f(\tau)$ be the Fourier series

$$f(\tau) = \sum_{n=1}^{\infty} \frac{a_n}{n^{k-1}} q^n.$$

Let

$$L(F,s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s}$$

and finally,

$$h(\tau) = f(\tau) - \sum_{0\le r<\frac{1}{2}(k-2)} \frac{L(F,k-r-1)}{k!}(2\pi i\tau)^r.$$

Then

$$h(\tau) - D = (-1)^{k-1}\varepsilon(-i\tau\sqrt{N})^{k-2} h\left(-\frac{1}{N\tau}\right)$$

where $D = 0$ if k is odd and $D = L(F,\frac{1}{2}k)(2\pi i\tau)^{\frac{1}{2}k-1}/(\frac{1}{2}k-1)!$ if k is even. Moreover, $L(F,\frac{1}{2}k) = 0$ if $\varepsilon = -1$.

Proof. We apply a lemma of Hecke, see [9, Section 5] with $G(\tau) = \varepsilon F(\tau)/(i\sqrt{N})^k$ to obtain

$$f(\tau) - \varepsilon(-1)^{k-1}(-i\tau\sqrt{N})^{k-2} f\left(-\frac{1}{N\tau}\right) = \sum_{r=0}^{k-2} \frac{L(F,k-r-1)}{r!}(2\pi i\tau)^r.$$

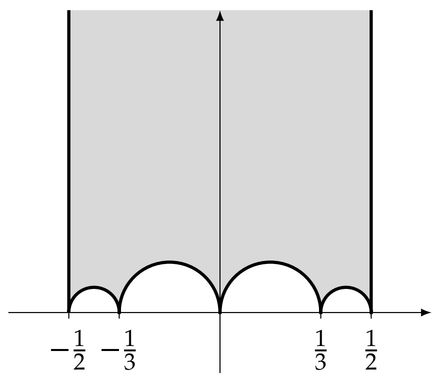
Split the summation on the right hand side into summations over $r < \frac{1}{2}k - 1$, $r > \frac{1}{2}k - 1$ and, possibly, $r = \frac{1}{2}k - 1$. For the region $r > \frac{1}{2}k - 1$ we apply the functional equation

$$\frac{L(F, k-r-1)}{r!} = \varepsilon(-1)^k(-i\sqrt{N})^{k-2}\left(-\frac{1}{N}\right)^{k-r-2}(2\pi i)^{k-2r-2}\frac{L(F, r+1)}{(k-r-2)!}$$

and substitute r by $k-2-r$. □

2 The group $\Gamma_1(6)$

This group is exactly the subgroup of $\mathrm{SL}_2(\mathbb{Z})$ of all matrices $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ with $a \equiv d \equiv 1 \pmod 6$, $c \equiv 0 \pmod 6$. Its fundamental domain can be pictured as below.



A complete set of inequivalent cusps is given by $0, 1/2, 1/3, \infty$. They are regular and have widths $6, 3, 2, 1$ respectively. Consider the following function

$$y(\tau) = \frac{\eta(6\tau)^8\eta(\tau)^4}{\eta(2\tau)^6\eta(3\tau)^4}$$

where

$$\eta(\tau) = q^{1/24}\prod_{n=1}^{\infty}(1-q^n), \quad q = e^{2\pi i\tau}, \quad \tau \in \mathbb{H}.$$

That it is a modular function on $\Gamma_1(6)$ can be checked using the tranformation formula for $\eta(\tau)$ in [6, Ch 9]. Since $y(\tau)$ has only one simple zero in the fundamental domain it generates the field of modular functions on $\Gamma_1(6)$. Moreover, $y(0) = \frac{1}{9}$, $y(\frac{1}{3}) = 1$, $y(\frac{1}{2}) = \infty$, $y(\infty) = 0$. The function $y(-\frac{1}{6\tau})$ is again invariant on $\Gamma_1(6)$ and one easily checks that

$$y\left(-\frac{1}{6\tau}\right) = \frac{y(\tau) - 1/9}{y(\tau) - 1}. \tag{1}$$

Hence the function

$$t(\tau) = y(\tau)\frac{1-9y(\tau)}{1-y(\tau)}$$

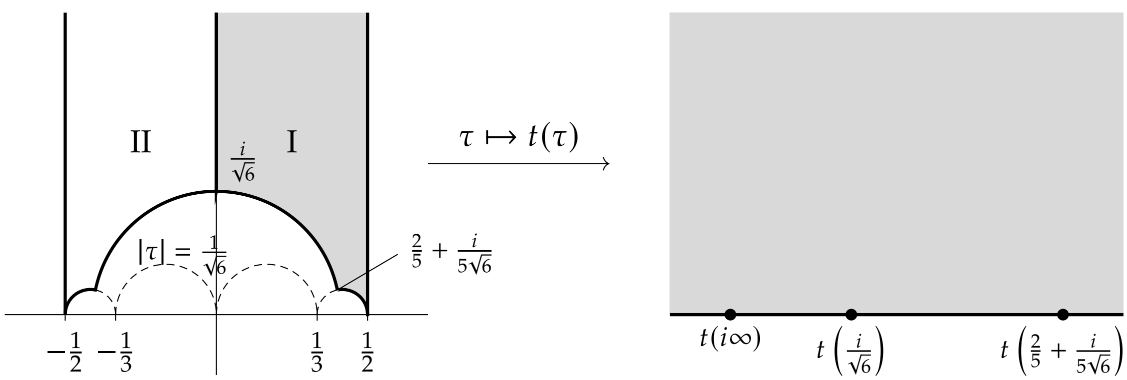
is invariant under the involution $\tau \mapsto -1/6\tau$. Moreover,

$$t(\tau) = \left(\frac{\Delta(6\tau)\Delta(\tau)}{\Delta(3\tau)\Delta(2\tau)}\right)^{1/2} = q\prod_{n=0}^{\infty}(1-q^{6n+1})^{12}(1-q^{6n+5})^{-12}$$

which is checked by noticing that $(\Delta(6\tau)\Delta(\tau)/\Delta(3\tau)\Delta(2\tau))^{1/2}$ is modular with respect to $\Gamma_1(6)$, invariant under $\tau \mapsto -1/6\tau$ and its zeros and poles coincide with those of $t(\tau)$.

Proposition 2.1. The function $t(\tau)$ maps the shaded open area in the picture below univalently onto the upper half plane and satisfies

$$t(i\infty) = 0, \quad t\left(\frac{i}{\sqrt{6}}\right) = (\sqrt{2}-1)^4, \quad t\left(\frac{2}{5}+\frac{i}{5\sqrt{6}}\right) = (\sqrt{2}+1)^4, \quad t\left(\frac{1}{2}\right) = \infty.$$



Proof. That $t(i\infty) = 0$, $t(\frac{1}{2}) = \infty$ can be ssen from the values $y(i\infty) = 0$, $y(\frac{1}{2}) = \infty$. From (1) it follows that for $\tau = \frac{i}{\sqrt{6}}$ and $y_0 = y(\frac{i}{\sqrt{6}})$, we have $y_0 = \frac{y_0-\frac{1}{9}}{y_0-1}$, hence $y_0 = 1 \pm \frac{2\sqrt{2}}{3}$ and correspondingly, $t(\frac{i}{\sqrt{6}}) = (\sqrt{2}\pm 1)^4$. The same principle can be applied to obatin $t(\frac{2}{5}+\frac{i}{5\sqrt{6}}) = (\sqrt{2}\pm 1)^4$. To decide which sign should be taken, one estimates $t(\frac{i}{\sqrt{6}})$ and $t(\frac{2}{5}+\frac{i}{5\sqrt{6}})$ numerically and obtain the values of our proposition. Furthermore, $t(\tau)$ assumes every value at most once in the union of I and II. Our proposition now follows. □

In the theorems and proofs that follow we let $\mathrm{M}_k(\Gamma_1(6))$ be the space of modular forms of weight k with respect to $\Gamma_1(6)$, and let

$$E_4(\tau) = 1 + 240\sum_{n=1}^{\infty}\sigma_3(n)q^n, \quad E_2(\tau) = 1 - 24\sum_{n=1}^{\infty}\sigma_1(n)q^n$$

be the standard Eisenstein series.

The group $\Gamma_1(6)$

Theorem 1. $\zeta(3)$ is irrational.

Proof. Let

$$40F(\tau) = E_4(\tau) - 36E_4(\tau) - 7(4E_4(2\tau) - 9E_4(3\tau))$$
$$24E(\tau) = -5(E_2(\tau) - 6E_2(6\tau)) + 2E_2(2\tau) - 3E_2(3\tau).$$

Notice that $F(\tau) \in \mathrm{M}_4(\Gamma_1(6))$ and $F(-\frac{1}{6\tau}) = -36\tau^4 F(\tau)$, $F(i\infty) = 0$ and $E(\tau) \in \mathrm{M}_2(\Gamma_1(6))$, $E(-\frac{1}{6\tau}) = -6\tau^2 E(\tau)$. The Dirichlet series corresponding to $F(\tau)$ reads

$$\begin{aligned} L(F,s) &= \sum_{n=1}^{\infty} \frac{6\sigma_3(n)}{n^s} - 36\frac{6\sigma_3(n)}{(6n)^s} - 28\frac{6\sigma_3(n)}{(2n)^s} + 63\frac{6\sigma_3(n)}{(3n)^s} \\ &= 6(1 - 6^{2-s} - 7\cdot 2^{2-s} + 7\cdot 3^{2-s})\zeta(s)\zeta(s-3). \end{aligned}$$

Define $f(\tau)$ by $(\frac{\mathrm{d}}{\mathrm{d}\tau})^3 f(\tau) = (2\pi i)^3 F(\tau)$, $f(i\infty) = 0$. From Proposition 1.2 and the fact that $F(-\frac{1}{6\tau}) = -36\tau^4 F(\tau)$ follows[2]

$$6\tau^2\left(f\left(-\frac{1}{6\tau}\right) - L(F,3)\right) = -(f(\tau) - L(F,3))$$

and since $L(F,3) = 6\cdot(-1/3)\zeta(3)\zeta(0) = \zeta(3)$, we have

$$6\tau^2\left(f\left(-\frac{1}{6\tau}\right) - \zeta(3)\right) = -(f(\tau) - \zeta(3)).$$

Multiplication with $E(-1/6\tau) = -6\tau^2 E(\tau)$ gives

$$E\left(-\frac{1}{6\tau}\right)\left(f\left(-\frac{1}{6\tau}\right) - \zeta(3)\right) = E(\tau)(f(\tau) - \zeta(3)). \tag{2}$$

The function $E(\tau)(f(\tau) - \zeta(3))$ can be considered as a multivalued function of $t = t(\tau)$. We choose it at $t = 0$ as follows. From the expansion $t = q\prod_{n=1}^{\infty}(1 - q^{6n+1})^{12}(1 - q^{6n+5})^{-12} = q - 12q^2 + 66q^3 - 220q^4 + 495q^5 - \ldots$ one infers the inverse expansion $q = t + 12t^2 + 222t^3 + \ldots$. Then, from $E(\tau) = 1 + 5q + 13q^2 + \ldots$ one finds $E(t) = 1 + 5t + 73t^2 + 1445t^3 + \ldots$ and similarly, $E(t)F(t) = 6t + (351/4)t^2 + \ldots$.

By construction one notes that $E(t) \in \mathbb{Z}[\![t]\!]$ and and $E(t)f(t) = \sum_{n=1}^{\infty} a_n t^n$ where $a_n \in \mathbb{Z}/[1,\ldots,n]^3$. Since the inverse function $t \to \tau$ branches at $(\sqrt{2}-1)^4$ one expects the radius of convergence of $E(t)(f(t)-\zeta(3))$ to be $(\sqrt{2}-1)^4$. However, by the property (2), the function $t \mapsto E(t)(f(t) - \zeta(3))$ has no branch point at $t = (\sqrt{2}-1)^4$, and its radius of convergence equals at least the next branching value, which is $(\sqrt{2}+1)^4$. Furthermore, it cannot be a polynomial in t, since then $f(\tau) - \zeta(3)$ would be a modular form of weight -2, which is impossible. We now apply Proposition 1.1 with $\theta_1 = \zeta(3)$ to conclude $\zeta(3) \notin \mathbb{Q}$. □

[2] There's a typo in the original article: $L(F,s)$ should be $L(F,3)$.

The group $\Gamma_1(6)$

Remark. Note that $1, 5, 73, 1445, \ldots$ are exactly Apéry's numbers for $\zeta(3)$.

Theorem 2. Let $F(\tau) = \eta(\tau)^2\eta(2\tau)^2\eta(3\tau)^2\eta(6\tau)^2$ and $L(F, s)$ the corresponding Dirichlet series. Then at least one of the numbers $\pi^{-2}L(F, 2)$ and $L(F, 3) + \frac{47L(F,2)\zeta(3)}{48\pi^2}$ is irrational.

Proof. The function $F(\tau)$ is in $\mathrm{M}_4(\Gamma_1(6))$, it is a cusp form, and $F(-\frac{1}{6\tau}) = 36\tau^4F(\tau)$. Let $f(\tau)$ be the Fourier series such that $(\frac{\mathrm{d}}{\mathrm{d}\tau})^3 f(\tau) = (2\pi i)^3F(\tau)$, $f(i\infty) = 0$. Then it follows from Proposition 1.2 that

$$6\tau^2\left(f\left(-\frac{1}{6\tau}\right) - L(F, 3)\right) = f(\tau) - L(F, 3) - L(F, 2)(2\pi i\tau). \tag{3}$$

Consider also

$$240G(\tau) = 13(E_4(\tau) + 36E_4(6\tau)) - 37(4E_4(2\tau) + 9E_4(3\tau)).$$

It ahs the properties $G(i\infty) = 0$, $G(-\frac{1}{6\tau}) = 36\tau^4G(\tau)$. The corresponding Dirichlet series reads

$$L(G, s) = (13 + 13 \cdot 6^{2-s} - 37 \cdot 2^{2-s} - 37 \cdot 3^{2-s})\zeta(s)\zeta(s - 3).$$

Letting $(\frac{\mathrm{d}}{\mathrm{d}\tau})^3 g(\tau) = (2\pi i)^3G(\tau)$, $g(i\infty) = 0$, we have

$$6\tau^2\left(g\left(-\frac{1}{6\tau}\right) - L(G, 3)\right) = g(\tau) - L(G, 3) - L(G, 2)(2\pi i\tau),$$

hence,

$$6\tau^2\left(g\left(-\frac{1}{6\tau}\right) - \frac{47}{3}\zeta(3)\right) = g(\tau) - \frac{47}{6}\zeta(3) + 48\zeta(2)(2\pi i\tau). \tag{4}$$

Elimination of $2\pi i\tau$ form (3) and (4) gives that the function $h(\tau) = 48\zeta(2)(f(\tau) - L(F, 3)) + L(F, 2)(g(\tau) - \frac{47}{6}\zeta(3))$ behaves like $6\tau^2h(-\frac{1}{6\tau}) = h(\tau)$. Now consider

$$E(\tau) = E_2(\tau) - 2E_2(2\tau) + 6E_2(3\tau) - 3E_2(6\tau).$$

It is in $\mathrm{M}_2(\Gamma_1(6))$ and we have $E(-\frac{1}{6\tau}) = 6\tau^2E(\tau)$. Consequently, $E(-\frac{1}{6\tau})h(-\frac{1}{6\tau}) = E(\tau)h(\tau)$ and by an argument to the one in Theorem 1 we find that

$$48\zeta(2)f(t)E(t) + L(F, 2)g(t)E(t) - \left(48\zeta(2)L(F, 3) + L(F, 2)\frac{47}{6}\zeta(3)\right)E(t)$$

is a power series in t with radius of convergence $(\sqrt{2}+1)^4$. Again the denominator of the n-th coefficient in the power series of $E(t)f(t)$, $E(t)g(t)$, $E(t)$ divides $[1, \ldots, n]^3$. We can now apply Proposition 1.1 to obtain our theorem. □

Theorem 3. Let

$$\sum_{n=1}^{\infty} a_n q^n = \left(\frac{\eta^9(\tau)\eta^9(6\tau)}{\eta^3(2\tau)\eta^3(3\tau)}\right)^{\frac{1}{2}} = q\prod_{n=1}^{\infty}(1-q^n)^3(1-q^{3n})^3\frac{(1+q^{3n})^{\frac{9}{2}}}{(1+q^n)^{\frac{3}{2}}}.$$

Then $\sum_{n=1}^{\infty} a_n/n^2$ is irrational.

Proof. Consider the product

$$E(\tau) = \frac{\eta^7(2\tau)\eta^7(3\tau)}{\eta^5(\tau)\eta^5(6\tau)}.$$

We have $E(-\frac{1}{6\tau}) = -6\tau^2 E(\tau)$ and hence $\sqrt{E(-\frac{1}{6\tau})} = \pm(i\tau\sqrt{6})\sqrt{E(\tau)}$. Since $E(\tau)$ has only zeros and poles in the cusps, it can be well-defined on the upper half plane. Since $E(\frac{i}{\sqrt{6}}) \neq 0$, we should have $\sqrt{E(-\frac{1}{6\tau})} = -i\tau\sqrt{6}E(\tau)$. Now consider

$$F(\tau) = \frac{\eta^7(\tau)\eta^7(6\tau)}{\eta^5(2\tau)\eta^5(3\tau)}\sqrt{E(\tau)}$$

which obeys $F(-\frac{1}{6\tau}) = (-i\tau\sqrt{6})^3 F(\tau)$. Let $f(\tau)$ be defined by $(\frac{\mathrm{d}}{\mathrm{d}\tau})^2 f(\tau) = (2\pi i)^2 F(\tau)$, $f(i\infty) = 0$. Then

$$-i\tau\sqrt{6}\left(f\left(-\frac{1}{6\tau}\right) - L(F,2)\right) = f(\tau) - L(F,2).$$

Multiplication with $\sqrt{E(-\frac{1}{6\tau})} = -i\tau\sqrt{6}\sqrt{E(\tau)}$ yields[3]

$$\sqrt{E\left(-\frac{1}{6\tau}\right)}\left(f\left(-\frac{1}{6\tau}\right) - L(F,2)\right) = \sqrt{E(\tau)}(f(\tau) - L(F,2)).$$

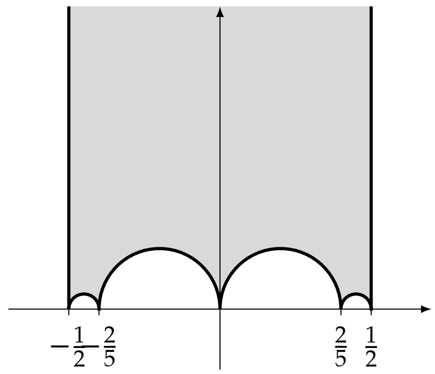
Notice that $\sqrt{E(\tau)}$ considered as a function of t is a power series whose n-th coefficient is rational and has denominator dividing $4^n[1,\ldots,n]^2$. By the same argument as in the previous theorems, the radius of convergence of $\sqrt{E(t)}(f(t) - L(F,2))$ is at least $(\sqrt{2}+1)^4$. Since $4e < (\sqrt{2}+1)^4$, we can apply Proposition 1.1 to find our theorem. □

Remark. Theorem 3 is the one alluded to in [1].

[3] There's a typo in the original article: $f(\tau)$ on the LHS should be $f(-\frac{1}{6\tau})$.

3 The group $\Gamma_1(5)$

The fundamental domain of the group $\Gamma_1(5) = \{(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}) \in SL_2(\mathbb{Z}) : a \equiv d \equiv 1 \pmod 5, c \equiv 0 \pmod 5\}$ can be pictured as below. The cusps are given by $0, \frac{1}{2}, \frac{2}{5}, i\infty$.



They are regular and have widths $5, 5, 1, 1$ respectively. Consider the following function,

$$y(\tau) = q \prod_{n=1}^{\infty} (1 - q^n)^{5\left(\frac{n}{5}\right)}$$

where $\left(\frac{n}{5}\right)$ it the Legendre symbol. The function $y(\tau)$ is a hauptmodul for the group $\Gamma_1(5)$. Moreover, $y(0) = -\frac{11}{2} + \frac{5}{2}\sqrt{5}$, $y(\frac{2}{5}) = i\infty$, $y(\frac{1}{2}) = -\frac{11}{2} - \frac{5}{2}\sqrt{5}$, $y(i\infty) = 0$. The function $y(-\frac{1}{5\tau})$ is again modular with respect to $\Gamma_1(5)$ and one easily checks that

$$y\left(-\frac{1}{5\tau}\right) = \frac{\lambda_1 - y(\tau)}{1 + \lambda_1 y(\tau)}, \quad \lambda_1 = -\frac{11}{2} + \frac{5}{2}\sqrt{5}.$$

So the function

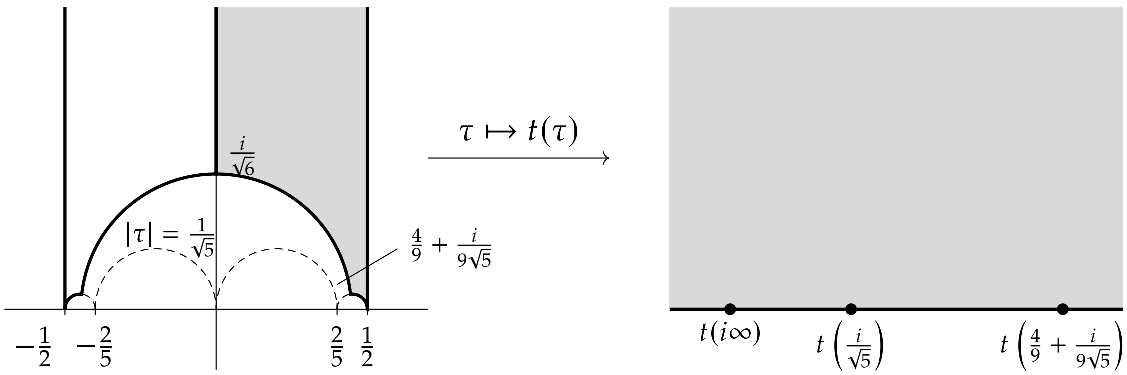
$$t(\tau) = y(\tau)\frac{\lambda_1 - y(\tau)}{1 + \lambda_1 y(\tau)}$$

is invariant under the involution $t \mapsto -1/5\tau$. In a similar way as in Proposition 2.1 one shows,

Proposition 3.1. The function $t(\tau)$ maps the shaded open area in the picture below univalently onto the upper half plane and satisfies

$$t(i\infty) = 0, \quad t\left(\frac{i}{\sqrt{5}}\right) = (\lambda_2 + \sqrt{1 + \lambda_2^2})^2, \quad t\left(\frac{4}{9} + \frac{i}{9\sqrt{5}}\right) = (\lambda_2 - \sqrt{1 + \lambda_2^2})^2, \quad t\left(\frac{1}{2}\right) = \infty$$

where $\lambda_2 = \frac{11}{2} - \frac{5}{2}\sqrt{5}$.



We also consider the function

$$s(\tau) = y(\tau)\frac{\lambda_2 - y(\tau)}{1 + \lambda_2 y(\tau)}, \quad \lambda_2 = \frac{11}{2} - \frac{5}{2}\sqrt{5}.$$

Lemma 3.2. The branching values of $s(\tau)$, as defined in Section 1, read $0, \infty$ and $(\lambda_1 \pm \sqrt{1+\lambda_1^2})^2$ where $\lambda_1 = -\frac{11}{2} + \frac{5}{2}\sqrt{5}$.

Proof. The branching values of $s(\tau)$ are the values of $s(\tau)$ at the cusps or the values at the points τ, $\Im\tau > 0$ where $s'(\tau) = 0$. The values at the cusps are $0, \infty$. Notice that

$$\frac{s'}{s} = \left(1 + \frac{y}{y - \lambda_2} - \frac{y}{y - \lambda_1}\right)\frac{y'}{y} = \frac{y^2 + (11 - 5\sqrt{5})y - 1}{y^2 + 11y - 1}\frac{y'}{y}.$$

The function $\frac{y'}{y}$ can only be zero at the cusps $0, \frac{1}{2}$. If $y^2 + (11 - 5\sqrt{5})y - 1 = 0$ then $y = \lambda_1 \pm \sqrt{1+\lambda_1^2}$ which implies $s = (\lambda_1 \pm \sqrt{1+\lambda_1^2})^2$. □

Notice that the q-expansions of $t(\tau)$, $s(\tau)$ read

$$t(\tau) = \sum_{n=1}^{\infty} a_n q^n, \quad s(\tau) = \sum_{n=1}^{\infty} b_n q^n$$

where a_n, b_n are algebraic integers in $\mathbb{Q}(\sqrt{5})$. From the construction follows that for every n the numbers a_n, b_n are conjugates.

Theorem 4. Let $L(3,\chi) = \sum_{n=1}^{\infty} \left(\frac{n}{5}\right) n^{-3}$, where $\left(\frac{n}{5}\right)$ is the Legendre symbol. Then $8\zeta(3) - 5\sqrt{5}L(3,\chi)$ is not in $\mathbb{Q}(\sqrt{5})$.

Proof. Consider the weight 4 form on $\Gamma_1(5)$ given by

$$24F(\tau) = E_4(\tau) - 25E_4(5\tau) + 24(E_4(\chi,\tau) - 5\sqrt{5}F_4(\chi,\tau))$$

where

$$E_4(\chi,\tau) = 1 + \sum_{n=1}^{\infty} \left(\frac{n}{5}\right) \frac{n^3 q^n}{1-q^n},$$
$$F_4(\chi,\tau) = \sum_{m,n=1}^{\infty} \left(\frac{n}{5}\right) m^3 q^{mn}.$$

Note that up to a constant factor, $F(\tau)$ is cahracterised by the facts $F(\tau) \in \mathrm{M}_4(\Gamma_1(5))$, $F(i\infty) = 0$, $F(-\frac{1}{5\tau}) = -25\tau^4 F(\tau)$. The corresponding Dirichlet series reads

$$L(F,s) = 10(1-5^{2-s})\zeta(s)\zeta(s-3) + \zeta(s)L(s-3,\chi) - 5\sqrt{5}\zeta(s-3)L(s,\chi)$$

where

$$L(s,\chi) = \sum_{n=1}^{\infty} \left(\frac{n}{5}\right) n^{-s}.$$

Define $f(\tau)$ by $f(i\infty) = 0$, $(\frac{\mathrm{d}}{\mathrm{d}\tau})^3 f(\tau) = (2\pi i)^3 F(\tau)$. then, from Proposition 1.2 follows that

$$5\tau^2 \left(f\left(-\frac{1}{5\tau}\right) - A\right) = -(f(\tau) - A)$$

where

$$\begin{aligned} A &= 10\left(1-\frac{1}{5}\right)\zeta(3)\zeta(0) + \zeta(3)L(0,\chi) - 5\sqrt{5}\zeta(0)L(3,\chi) \\ &= -\frac{1}{3}(8\zeta(3) - 5\sqrt{5}L(3,\chi)). \end{aligned}$$

Now let

$$-8E(\tau) = E_2(\tau) - 5E_2(5\tau) + 20(E_2(\chi,\tau) - \sqrt{5}F_2(\chi,\tau))$$

where

$$E_2(\chi,\tau) = -\frac{1}{5} + \sum_{n=1}^{\infty} \left(\frac{n}{5}\right) \frac{nq^n}{1-q^n},$$
$$F_2(\chi,\tau) = \sum_{m,n=1}^{\infty} \left(\frac{n}{5}\right) mq^{mn}.$$

The function $E(\tau)$ satisfies $E(-\frac{1}{5\tau}) = -5\tau^2 E(\tau)$, hence $E(\tau)(f(\tau) - A)$ is fixed under the involution $\tau \mapsto -1/5\tau$. Consider $E(\tau)$ and $E(\tau)f(\tau)$ as functions of $t = t(\tau)$ and write

$$E(\tau)f(\tau) = \sum_{n=1}^{\infty} c_n q^n, \quad E(\tau) = \sum_{n=1}^{\infty} d_n q^n.$$

By construction it follows that d_n and $[1,\dots,n]^3 c_n$ are algebraic integers in $\mathbb{Q}(\sqrt{5})$. Just as in the proof of Theorem 1, we observe that the radius of convergence of $E(t)(f(t)-A)$ equals $(\lambda_2+\sqrt{1+\lambda_2^2})^2$[4] and hence for all $\epsilon>0$,

$$|c_n - Ad_n| < (\lambda_2+\sqrt{1+\lambda_2^2})^{(2-\epsilon)n} \quad \forall n > n_0(\epsilon). \tag{5}$$

Now consider the functions

$$24\overline{F}(\tau) = E_4(\tau) - 25E_4(5\tau) + 24(E_4(\chi,\tau) + 5\sqrt{5}F_4(\chi,\tau))$$

the corresponding third primitive $\overline{f}(\tau)$ and

$$-8\overline{E}(\tau) = E_2(\tau) - 5E_2(5\tau) + 20(E_2(\chi,\tau) + \sqrt{5}F_2(\chi,\tau)).$$

Consider them as functions of $s = s(\tau)$ and write

$$\overline{E}(\tau)\overline{f}(\tau) = \sum_{n=1}^{\infty} \overline{c}_n q^n, \quad \overline{E}(\tau) = \sum_{n=1}^{\infty} \overline{d}_n q^n.$$

From the construction follows that $\overline{c}_n, \overline{d}_n$ are conjugates of c_n, d_n respectively. By Lemma 3.2 the smallest nonzero branching value of $s(\tau)$ equals $(-\lambda_1+\sqrt{1+\lambda_1^2})^2$ and hence the radius of convergence of both $\sum_{n=1}^{\infty}\overline{c}_n s^n$ and $\sum_{n=1}^{\infty}\overline{d}_n s^n$ is at least $(-\lambda_1+\sqrt{1+\lambda_1^2})^2$. Hence for any $\theta\in\mathbb{C}$ and any $\epsilon>0$

$$|\overline{c}_n - \theta\overline{d}_n| < (\lambda_1+\sqrt{1+\lambda_1^2})^{(2+\epsilon)n} \quad \forall n > n_1(\epsilon,\theta). \tag{6}$$

Now suppose $A\notin\mathbb{Q}(\sqrt{5})$. Let $\overline{A}$ be its conjugate and let d be its denominator. Multiplication of (5) and (6) with $\theta = \overline{A}$ yields

$$|c_n\overline{c}_n - (c_n\overline{d}_n\overline{A} + \overline{c}_n d_n A) + d_n\overline{d}_n A\overline{A}| < \left(\frac{1}{20.3}\right)^{2n} \quad \forall n > n_0. \tag{7}$$

Since $c_n\overline{c}_n \in \mathbb{Z}/[1,\dots,n]^6$, $c_n\overline{D}_n\overline{A}+\overline{c}_n d_n A \in \mathbb{Z}/d[1,\dots,n]^3$, $d_n\overline{d}_n A\overline{A}\in\mathbb{Z}/d^2$ and $d^2[1,\dots,n]^6 < (20.1)^{2n} < (20.3)^{2n}$ for sufficiently large n. Hence $c_n - d_n A = 0$ for n large enough, and we have a contradiction. Theorem 4 now follows. □

Remark. By some tedius calculation one can veritfy that the numbers d_n satisfy the recurrence relation

$$(n+1)^3 d_{n+1} = ((124+55\sqrt{5})n(n+1) + 34 + 15\sqrt{5})(2n+1)d_n - n^3 d_{n-1}$$
$$d_0 = 1, d_1 = 34+15\sqrt{5}, d_2 = 7111+3180\sqrt{5}, d_3 = 2040334+912465\sqrt{5}.$$

[4]There's a typo in the original article: sign is fixed.

The group $\Gamma_1(5)$

Theorem 5. The number $\zeta(2) = \frac{\pi^2}{6}$ is irrational.

Proof. Consider the function

$$\frac{2i}{5}F(\tau) = (2+i)E_3(\chi,\tau) - (2-i)E_3(\chi,\tau)$$

where

$$E_3(\chi,\tau) = \frac{-2+i}{5} + \sum_{i=1}^{\infty} \chi(k)\frac{k^2 q^k}{1-q^k}$$

and $\chi(k)$ is the odd character modulo 5 given by $\chi(2) = -i$ and $\overline{\chi}$ is its complex conjugate. Then $F(\tau) \in \mathrm{M}_3(\Gamma_1(5))$, and, in particular, $F(\frac{\tau}{5\tau+1}) = (5\tau+1)^3 F(\tau)$. Let $f(\tau)$ be the Fourier series determined by $f(i\infty) = 0$ and $(\frac{\mathrm{d}}{\mathrm{d}\tau})^2 f(\tau) = (2\pi i)^2 F(\tau)$. In a straightforward manner one can verify that

$$(5\tau+1)\left(f\left(\frac{\tau}{5\tau+1}\right) - L(F,2)\right) = f(\tau) - L(F,2)$$

where

$$L(F,2) = \frac{5}{2i}\zeta(2)((2+i)L(0,\chi) - (2-i)L(0,\overline{\chi})) = \zeta(2).$$

Consider also

$$E(\tau) = \frac{3+i}{2}E_1(\chi,\tau) + \frac{3-i}{2}E_1(\overline{\chi},\tau)$$

where

$$E_1(\chi,\tau) = \frac{3-i}{10} + \sum_{k=1}^{\infty} \chi(k)\frac{q^k}{1-q^k}.$$

Then $E(\tau) \in \mathrm{M}_1(\Gamma_1(5))$ and we obtain

$$E\left(\frac{\tau}{5\tau+1}\right)\left(f\left(\frac{\tau}{5\tau+1}\right) - \zeta(2)\right) = E(\tau)(f(\tau) - \zeta(2)).$$

This implies that $E(\tau)(f(\tau)-\zeta(2))$ considered as function of $y(\tau)$ does not branch above $y = -\frac{11}{2} + \frac{5}{2}\sqrt{5}$, corresponding to $\tau = 0$. Hence $E(\tau)(f(\tau) - \zeta(2))$ as a function of y is a Taylor series in y with radius of convergence $\frac{11}{2} + \frac{5}{2}\sqrt{5}$. Note that by construction $E(\tau)$ has a y-expansion with integral coefficients, and the nth coefficient in the y-expansion of $E(\tau)f(\tau)$ is rational with a denominator that divides $[1,\ldots,n]^2$. Out standard argument now yields $\zeta(2) \notin \mathbb{Q}$. □

Remark. Notice that $E(t) = 1+3t+19t^2+147t^3+\cdots$ and the numbers $1, 3, 19, 147, \ldots$ correspond exactly to Apéry's numbers for $\zeta(2)$. The function $E(\tau)$ is also discussed in [2, p59].

References

[1] Beukers, F. Irrationality of π^2, periods of an elliptic curve and $\Gamma_1(5)$. *Diophantine approximations and transcendental numbers (Luminy, 1982) 31* (1982), 47–66.

[2] Beukers, F. Some congruences for the Apéry numbers. *Journal of Number Theory 21*, 2 (1985), 141–155.

[3] Beukers, F. Another congruence for the Apéry numbers. *Journal of Number Theory 25*, 2 (1987), 201–210.

[4] Beukers, F., and Peters, C. A. M. A family of $K3$ surfaces and $\zeta(3)$. *J. Reine Angew. Math. 351* (1984), 42–54.

[5] Gessel, I. Some Congruences for Apery Numbers. *Journal of Number Theory 14*, 3 (1982), 362–368.

[6] Rademacher, H. *Topics in analytic number theory*, vol. 169. Springer Science & Business Media, 2012.

[7] Stienstra, J., and Beukers, F. On the Picard-Fuchs equation and the formal Brauer group of certain elliptic $K3$-surfaces. *Mathematische Annalen 271* (1985), 269–304.

[8] Van der Poorten, A. A proof that Euler missed...Apéry's proof of the irrationality of $\zeta(3)$. *Math. Intelligencer 1*, 4 (1979), 195–203.

[9] Weil, A. Remarks on Hecke's lemma and its use. In *Algebraic number theory (Kyoto Internat. Sympos., Res. Inst. Math. Sci., Univ. Kyoto, Kyoto, 1976)*. Japan Soc. Promotion Sci., Tokyo, 1977, pp. 267–274.