The sum of eight squares

Marco Morosin

26.02.2021

Theorem

Let $r_8(n) := \#\{(n_1, \ldots, n_8) \in \mathbb{Z}^8 \mid n_1^2 + \cdots + n_8^2 = n\}$; then

$$r_8(n) = 16 \sum_{1 \leqslant d | n} (-1)^{n-d} d^3$$

The main tool: θ series

$$\theta(t) := \sum_{n \in \mathbb{Z}} e^{-t\pi n^2} \quad (t \in \mathbb{R}_{>0})$$

$$\theta(-2iz) = \sum_{n \in \mathbb{Z}} e^{2z\pi i n^2} = \sum_{n \in \mathbb{Z}} q^{n^2} \quad (z \in \mathcal{H})$$

$$\theta(-2iz)^k = \sum_{(n_1, \ldots, n_k) \in \mathbb{Z}^k} q^{n_1^2 + \cdots + n_k^2} = \sum_{n \geqslant 0} r_k(n) q^n$$

Outline

- $\theta(-2iz)^8 \in \mathcal{M}_4(J)$ for some congruence subgroup J
- $\mathcal{M}_4(J)$ is generated by E_4, E_4^* with Fourier coefficients written in terms of $\sigma_3(n) := \sum_{1 \leqslant d | n} d^3$
- find the linear combination $\theta(-2iz)^8 = aE_4 + bE_4^*$
- equate coefficients to find $r_8(n)$ in terms of $\sigma_3(n)$

Let $\Gamma \leqslant \mathrm{SL}_2(\mathbb{Z})$, $\chi\colon \Gamma \to \mathbb{C}^\times$ be a character, $k \in \mathbb{Z}$.

Definition

$$\mathcal{M}_k(\Gamma, \chi) := \text{the set of } f \text{ such that}$$

- $f\colon \mathcal{H} \to \mathbb{C}$ is holomorphic,
- for all $\gamma \in \Gamma$, $f|_k\gamma = \chi(\gamma) f$,
- for all $\gamma \in \mathrm{SL}_2(\mathbb{Z})$, $\exists \lim (f|_k\gamma)(z) \in \mathbb{C}$ for $\operatorname{Im} z \to \infty$.

Clearly $\mathcal{M}_k(\Gamma) = \mathcal{M}_k(\Gamma, 1)$.

Definition

- define $\Theta(z) := \theta(-iz) = \sum_{n\in\mathbb{Z}} e^{\pi z i n^2}$, $z \in \mathcal{H}$
- define the *theta group* $J := \left\langle T^2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \, S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \right\rangle$

Theorem

There is a unique character $\chi\colon J = \langle T^2, S\rangle \to \mathbb{C}^\times$ *such that* $\chi(S) = -i$, $\chi(T^2) = 1$.
Moreover, $\Theta^2 \in \mathcal{M}_1(J, \chi)$.

Proof.

Recall:

- $\Theta(z) := \theta(-iz) = \sum_{n\in\mathbb{Z}} e^{\pi z i n^2}$
- for $\operatorname{Re} z > 0$, $\theta(1/z) = \sqrt{z}\theta(z)$

Therefore it is easy to verify:

- $\Theta^2|_1 T^2 = \Theta^2$, $\Theta^2|_1 S = -i\Theta^2$
- then $\chi(\gamma) :=$ eigenvalue of the action of γ on Θ^2 gives a well-defined group morphism by the properties of an action, and

$$\Theta^2|_1\gamma = \chi(\gamma)\Theta^2 \qquad (\gamma \in J)$$

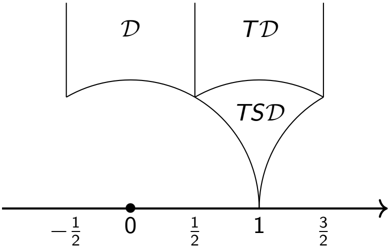
$\square$

Corollary

$\Theta^8 \in \mathcal{M}_4(J)$, *because* $\Theta^8|_4\gamma = \chi(\gamma)^4\Theta^8 = \Theta^8$.

Proposition

1. *for every $z \in \mathcal{H}$, there is $\gamma \in J$ such that $\gamma z \in \mathcal{D} \cup T\mathcal{D} \cup TS\mathcal{D}$*
2. $J\backslash \mathrm{SL}_2(\mathbb{Z}) = \{J, JT, JTS\}$*. In particular,* $[\mathrm{SL}_2(\mathbb{Z}) : J] = 3$



$$\begin{cases} -\frac{1}{2} \leqslant \operatorname{Re} z \leqslant \frac{3}{2} \\ |z| \geqslant 1 \\ |z-2| \geqslant 1 \end{cases}$$

Proof of 1.

Let $z \in \mathcal{H}$; we show $Jz \cap (\mathcal{D} \cup T\mathcal{D} \cup TS\mathcal{D}) \neq \varnothing$:

- take $z' \in Jz$ with maximal $\operatorname{Im} z'$ (recall $\operatorname{Im} \gamma z = \operatorname{Im} z/|cz+d|^2$)
- choose it with $-\frac{1}{2} \leqslant |\operatorname{Re} z'| \leqslant \frac{3}{2}$ (we can always translate by $T^2 \in J$)
- $-1/z' \in Jz$ and $\operatorname{Im}(-1/z') = \frac{\operatorname{Im} z'}{|z'|^2}$, so $|z'| \geqslant 1$
- $ST^{-2}z' = \left(\begin{smallmatrix} 0 & -1 \\ 1 & -2 \end{smallmatrix}\right) z' \in Jz$ and $\operatorname{Im}(ST^{-2}z') = \frac{\operatorname{Im} z'}{|z'-2|^2}$, so $|z'-2| \geqslant 1$
- hence $z' \in Jz \cap (\mathcal{D} \cup T\mathcal{D} \cup TS\mathcal{D})$.

Proof of 2. ($J\backslash\mathrm{SL}_2(\mathbb{Z}) = \{J, JT, JTS\}$)

- take $z' \in \mathring{\mathcal{D}}$: then for $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ we have ($\gamma z' \in \mathcal{D} \iff \gamma = \pm 1_2$)
- let $\gamma \in \mathrm{SL}_2(\mathbb{Z})$; take $\delta \in J$ such that $\delta\gamma z' \in \mathcal{D} \cup T\mathcal{D} \cup TS\mathcal{D}$ (by the previous point)
- then $\delta\gamma = \pm 1_2$ or $T^{-1}\delta\gamma = \pm 1_2$ or $(TS)^{-1}\delta\gamma = \pm 1_2$
  i.e. $\gamma = \pm\delta^{-1}1_2$ or $\gamma = \pm\delta^{-1}T$ or $\gamma = \pm\delta^{-1}TS$
- Moreover $-1_2 = S^2 \in J$, hence $\mathrm{SL}_2(\mathbb{Z}) = J \cup JT \cup JTS$
- we check the union is disjoint: consider $\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/2\mathbb{Z})$, $\gamma \mapsto \overline{\gamma}$ and check the action of $\overline{\gamma}^{-1}$ on $1 = \left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)$

$$\text{for } \gamma \in J,\ \overline{\gamma}^{-1}\colon 1 \mapsto 1$$

$$\text{for } \gamma = \delta T \in JT,\ \overline{\gamma}^{-1} = \overline{T}^{-1}\overline{\delta}^{-1}\colon 1 \mapsto 0$$

$$\text{for } \gamma = \delta TS \in JTS,\ \overline{\gamma}^{-1} = (\overline{TS})^{-1}\overline{\delta}^{-1}\colon 1 \mapsto \infty$$

$\square$

Note, in particular, we showed $J \supset \Gamma(2)$.

Proposition

Let $\Gamma \leqslant \mathrm{SL}_2(\mathbb{Z})$ *of finite index* $[\mathrm{SL}_2(\mathbb{Z}) : \Gamma] = m$.
Let χ *be a character with* $\chi^e = 1$ *for some* $e \in \mathbb{Z}$*. Then, for* $k \geqslant 0$,

$$\dim \mathcal{M}_k(\Gamma, \chi) \leqslant \frac{mk}{12} + 1.$$

- note: since $\chi^e = 1$, then $f \in \mathcal{M}_k(\Gamma, \chi) \implies f^e \in \mathcal{M}_{ek}(\Gamma)$

for $f \in \mathcal{M}_k(\Gamma)$, we can define

$$N(f) := \prod_{[\gamma] \in \Gamma \backslash \mathrm{SL}_2(\mathbb{Z})} f|_k \gamma$$

- $N(f) \in \mathcal{M}_{mk}(\mathrm{SL}_2(\mathbb{Z}))$ since multiplication by γ is bijective on $\Gamma \backslash \mathrm{SL}_2(\mathbb{Z})$
- $N(f) = 0 \iff f = 0$: indeed, if $N(f) = 0$, then $f|_k\gamma = 0$ for some γ (otherwise $N(f)$ would only have isolated zeros), so $f = (f|_k\gamma)|_k\gamma^{-1} = 0$

Proof of the Proposition

Let $P \subset \mathring{\mathcal{D}}$ with $\#P = \frac{mk}{12} + 1 =: N$.
It is enough to show injectivity of the linear map

$$\psi \colon \mathcal{M}_k(\Gamma, \chi) \to \mathbb{C}^N, \quad f \mapsto (f(p))_{p \in P}$$

- let $f \in \ker \psi$; write $N(f^e) = f^e h$ with h holomorphic
- $v_p(N(f^e)) \geqslant e$ for every $p \in P$
- $N(f^e) \in \mathcal{M}_{emk}(\mathrm{SL}_2(\mathbb{Z}))$. If it's non-zero, we can apply the $emk/12$-formula to $N(f^e)$:

$$\frac{emk}{12} = v_\infty(N(f^e)) + \frac{v_i(N(f^e))}{2} + \frac{v_\rho(N(f^e))}{3} + \sum_{\substack{[z] \in \mathrm{SL}_2(\mathbb{Z}) \backslash \mathcal{H} \\ z \neq i, \rho}} v_z(N(f^e))$$

- but $\sum_{\substack{[z] \\ z \neq i, \rho}} v_z(N(f^e)) \geqslant Ne = (\frac{mk}{12} + 1)e > \frac{emk}{12}$: contradiction.
- Hence, $N(f^e) = 0 \implies f^e = 0 \implies f = 0$

It follows that $\dim_{\mathbb{C}} \mathcal{M}_k(\Gamma, \chi) \leqslant N$.

$\square$

Application to $M_4(J)$

Recall $[\mathrm{SL}_2(\mathbb{Z})\colon J] = 3$, hence

$$\dim \mathcal{M}_4(J) \leqslant \frac{3 \cdot 4}{12} + 1 = 2.$$

Proposition

Let $k \geqslant 4$ even. The following series converges absolutely on $\mathcal{H}$ and defines an element of $\mathcal{M}_k(J)$.

$$G_k^*(z) := \sum_{\substack{(m,n)\neq(0,0) \\ m\equiv n \text{ mod } 2}} \frac{1}{(mz+n)^k}$$

Proof

Recall $G_k(z) = \sum_{(m,n)\neq(0,0)} \frac{1}{(mz+n)^k}$ converges absolutely on $\mathcal{H}$; hence also $G_k^*(z)$.

$$\begin{aligned} G_k^*(T^2 z) = G_k^*(z+2) = \textstyle\sum_{m\equiv n \bmod 2} (mz+2m+n)^{-k} &= \\ = \textstyle\sum_{m\equiv n \bmod 2} (mz+n)^{-k} &= G_k^*(z) \end{aligned}$$

because $(m,n) \mapsto (m, 2m+n)$ is a bijection of $\{(m,n) : m \equiv n \bmod 2\}$

$$\begin{aligned} G_k^*(Sz) = G_k^*\left(-\frac{1}{z}\right) = \textstyle\sum_{m\equiv n \bmod 2} (-\frac{m}{z}+n)^{-k} = \sum_{m\equiv n \bmod 2} (\frac{-m+nz}{z})^{-k} &= \\ = z^k \textstyle\sum_{m\equiv n \bmod 2} (-m+nz)^{-k} &= z^k G_k^*(z) \end{aligned}$$

We need to check $\exists \lim_{y\to\infty}(G_k^*|_k\gamma)(z) \in \mathbb{C}$ for all $\gamma \in \mathrm{SL}_2(\mathbb{Z})$.

- By the transformation property, we only need to check for 1, T, TS.
- $(G_k^*|_k T)(z) = G_k^*(z+1)$ has finite limit $\iff$ $G_k^*(z)$ does, so it remains to check for 1 and TS

$$TS = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}$$

So, all we have to check is the existence of finite

$$\lim_{y\to\infty} G_k^*(z) =: G_k^*(\infty)$$

$$\lim_{y\to\infty} (G_k^*|_k TS)(z) = \lim_{y\to\infty} z^{-k} G_k^*\left(1 - \frac{1}{z}\right) =: G_k^*(1)$$

$G_k^*(\infty) = 2^{1-k}\zeta(k)$

- $G_k^*(z) = \sum_{\substack{(m,n)\neq(0,0)\\ m\equiv n \bmod 2}} (mz+n)^{-k} = \sum_{(m,n)\neq(0,0)} (mz+m+2n)^{-k}$
- by uniform convergence, exchange lim and $\sum$:

$$\lim G_k^*(z) = \lim \sum_{(m,n)\neq(0,0)} (mz+m+2n)^{-k} = \sum_{(m,n)\neq(0,0)} \lim\,(mz+m+2n)^{-k}$$

- $\lim_{y\to\infty} (mz+m+2n)^{-k} = 0$ if $m \neq 0$; hence

$$\cdots = \sum_{n\neq 0} (2n)^{-k} = 2\sum_{n\geqslant 1} 2^{-k} n^{-k} = 2^{1-k}\zeta(k)$$

$G_k^*(1) = 2\zeta(k)$

$$z^{-k}G_k^*(1-1/z) = z^{-k}\sum_{m\equiv n}(m(1-1/z)+n)^{-k} = z^{-k}\sum_{m\equiv n}\left(\frac{mz-m+nz}{z}\right)^{-k}$$
$$= \sum_{m\equiv n}(mz-m+nz)^{-k} = \sum_{(m,n)}(2mz+n)^{-k}$$

$$\lim_{y\to\infty} z^{-k}G_k^*(1-1/z) = \sum_{(m,n)} \lim_{y\to\infty}(2mz+n)^{-k} = \sum_{(0,n)} n^{-k} = 2\sum_{n\geqslant 1} n^{-k} = 2\zeta(k)$$

$\square$

Then $E_k^*(z) := \frac{2^{k-1}}{\zeta(k)} G_k^*(z) \in \mathcal{M}_k(J)$ satisfies $E_k^*(\infty) = 1,\ E_k^*(1) = 2^k$.

Fourier expansion for E_k^*

Note that

$$G_k\left(\frac{z+1}{2}\right) = \sum_{(m,n)\neq(0,0)} \left(\frac{m}{2}z + \frac{m}{2} + n\right)^{-k} = \sum_{(m,n)\neq(0,0)} \left(\frac{mz+m+2n}{2}\right)^{-k} =$$

$$= 2^k \sum_{(m,n)\neq(0,0)} (mz+m+2n)^{-k} = 2^k G_k^*(z)$$

because the set of $(m, m+2n)$ is precisely the set $(m,n) : m \equiv n \bmod 2$.

Recalling $E_k(z) = 1 - \frac{2k}{B_k} \sum_{n\geqslant 1} \sigma_{k-1}(n) q^n$, we get

$$E_k^*(z) := \frac{2^{k-1}}{\zeta(k)} G_k^*(z) = \frac{2^{-1}}{\zeta(k)} G_k\left(\frac{z+1}{2}\right) =$$

$$= E_k\left(\frac{z+1}{2}\right) = 1 - \frac{2k}{B_k} \sum_{n\geqslant 1} (-1)^n \sigma_{k-1}(n) e^{i\pi nz}$$

Proposition

$\mathcal{M}_4(J) = \langle E_4, E_4^* \rangle$.

Proof.

Compute as before $G_k(\infty) = 2\zeta(k) = G_k(1) \implies E_k(\infty) = 1 = E_k(1)$.
Consider the linear map

$$u\colon \mathcal{M}_4(J) \to \mathbb{C}^2, \quad f \mapsto (f(\infty), f(1))$$

- $E_4, E_4^* \in \mathcal{M}_4(J)$ and are linearly independent since $u(E_4^*) = (1, 2^4)$ and $u(E_4) = (1, 1)$
- recalling $\dim \mathcal{M}_4(J) \leqslant \frac{4\cdot 3}{12} + 1 = 2$, we conclude.

$\square$

Remember $\Theta^8 \in \mathcal{M}_4(J)$; hence, $\Theta^8 = aE_4 + bE_4^*$ for some $a, b \in \mathbb{C}$.

Note: $\Theta^2(\infty) = 1$, since $\Theta(z) = \sum_{n\in\mathbb{Z}} e^{\pi i z n^2} = \sum_{n\in\mathbb{Z}} e^{\pi i x n^2} e^{-\pi y n^2}$.

Proposition

$\Theta^2(1) := \lim_{y\to\infty} z^{-1}\Theta^2(1-\frac{1}{z}) = 0$

Proof

We prove $\frac{1}{\sqrt{-iz}}\Theta(1-\frac{1}{z}) = \sum_{n\in\mathbb{Z}} e^{i\pi z(n+\frac{1}{2})^2}$, which will imply that

$$z^{-1}\Theta^2\left(1-\frac{1}{z}\right) = -i\left(\sum_{n\in\mathbb{Z}} e^{i\pi z\left(n+\frac{1}{2}\right)^2}\right)^2 \to 0$$

We consider $(\psi : x \mapsto e^{i\pi z x^2 + 2i\pi w x}) \in \mathcal{S}(\mathbb{R})$, $w\in\mathbb{C}$, $z\in\mathcal{H}$.
We first show $\widehat{\psi}(y) = \frac{1}{\sqrt{-iz}} e^{-i\frac{\pi}{z}(x-w)^2}$:

- set $f_z(x) = e^{i\pi z x^2}$, then $\widehat{\psi}(y) = \widehat{f_z}(y-w)$ so we may assume $w=0$
- the equality to prove becomes $\widehat{f_z}(y) = \frac{1}{\sqrt{-iz}} f_{-\frac{1}{z}}(y)$
- we know it holds on $z = it$ $(t\in\mathbb{R}_{>0})$ (Lemma 0.15)
- we conclude by noting that both sides are holomorphic functions of $z\in\mathcal{H}$

Now Poisson's formula $\sum_{n\in\mathbb{Z}}\psi(n) = \sum_{n\in\mathbb{Z}}\widehat{\psi}(n)$ with $w = z/2$ gives the result.

$\square$

Theorem

Let $r_8(n) := |\{(n_1, \dots, n_8) \in \mathbb{Z}^8 \mid n_1^2 + \cdots + n_8^2 = n\}|$*; then*

$$r_8(n) = 16 \sum_{1 \leqslant d | n} (-1)^{n-d} d^3$$

Proof

- recall $u(\Theta^8) = (1, 0),\ u(E_4) = (1, 1),\ u(E_4^*) = (1, 16)$
- write $\Theta^8 = aE_4 + bE_4^*$, hence $u(\Theta^8) = (1, 0) = (a + b, a + 16b)$
- we get $\Theta^8 = \frac{16}{15} E_4 - \frac{1}{15} E_4^*$

$$\sum_{n \geqslant 0} r_8(n) q^n = \Theta^8(2z) = \frac{1}{15}\Big(16 E_4(2z) - E_4^*(2z)\Big) =$$

$$= \frac{1}{15}\Big(15 + 16 \cdot 240 \sum_{n \geqslant 1} \sigma_3(n) q^{2n} - 240 \sum_{n \geqslant 1} (-1)^n \sigma_3(n) q^n\Big) =$$

$$= 1 + 16 \sum_{n \geqslant 1} \Big(16 \sigma_3(n) q^{2n} - (-1)^n \sigma_3(n) q^n\Big)$$

$\sum_{n\geqslant 0} r_8(n)q^n = 1 + \sum_{n\geqslant 1} 16\Big(16\sigma_3(n)q^{2n} - (-1)^n\sigma_3(n)q^n\Big)$

n odd: $r_8(n) = 16\sigma_3(n) = 16\sum_{d|n} d^3$
n even:

- $r_8(n) = 16(16\sigma_3(\frac{n}{2}) - \sigma_3(n))$
- $16\sigma_3(n/2) = 2\cdot 8\sum_{e|\frac{n}{2}} e^3 = 2\cdot\sum_{2e|n}(2e)^3 = 2\cdot\sum_{2|d|n} d^3$, hence

$$\begin{aligned} 16\sigma_3\left(\frac{n}{2}\right) - \sigma_3(n) &= 2\sum_{2|d|n} d^3 - \sigma_3(n) = \\ &= 2\sum_{2|d|n} d^3 - \sum_{d|n} d^3 = \sum_{d|n}(-1)^d d^3 \end{aligned}$$

therefore $r_8(n) = 16\sum_{d|n}(-1)^d d^3$

In any case: $r_8(n) = 16\sum_{d|n}(-1)^{n-d} d^3$ because

$$(-1)^{n-d} = \begin{cases} 1 & \text{if } n \text{ is odd (and hence } d \text{ is odd)} \\ (-1)^d & \text{if } n \text{ is even} \end{cases}$$

$\square$