Sums of Four and Eight Squares

Gilles Felber - Francesco Fournier Facio

May 14, 2019

1 Introduction and motivation

The goal of this talk is to prove explicit formulas to calculate, for a given positive integer n, the number of representations of n as a sum of k integers squares. Let us be more precise:

Definition 1.1. For positive integers k, n, we denote by

$$A_k(n) := \#\{(x_1, \dots, x_k) \in \mathbb{Z}^k : x_1^2 + \dots + x_k^k = n\}$$

the number of representations of n as a sum of k squared integers. Notice that the order, as well as the sign of the x_i is taken into account.

We will prove:

$$A_4(n) = 8 \sum_{4 \nmid d | n} d;$$

$$A_8(n) = 16 \sum_{d | n} (-1)^{n-d} d^3;$$

which are known as Jacobi's four- and eight-square formulas. Here both sums run over positive divisors of n. In the end, we will quickly mention other formulas for $A_k(n)$, when $k \neq 4, 8$.

But why should one care about such formulas? The source of the largest number of motivating examples lies in the following simple geometric interpretation: $A_k(n)$ is the number of points of $\mathbb{Z}^k$ intersecting the sphere of radius $\sqrt{n}$ in $\mathbb{R}^k$. This also motivates why we may wish to take into account also differences of order and/or of sign. These quantities appear frequently not only in geometry and number theory, but also in physics and crystallography.

One classical problem in number theory is what is commonly referred to as Gauss's circle problem. The question is to give an estimate of the number of points of $\mathbb{Z}^k$ that intersect the closed ball of radius $\sqrt{x}$ in $\mathbb{R}^k$. Using our notation, this quantity is equal to $\sum_{n \leq x} A_k(n)$. By a simple geometric argument, it is easy to see that

$$\sum_{n \leq x} A_2(n) = \pi x + O(\sqrt{x}).$$

Just associate to each point of $\mathbb{Z}^2$ in the disk a square of side-length 1, approximate the number of such squares by the area of the circle, and the error will be at most proportional to the circumference. More generally, with a similar argument:

$$\sum_{n\leq x} A_k(n) = \rho_k x^{\frac{k}{2}} + O(x^{\frac{k-1}{2}}),$$

where ρ_k is the volume of the unit ball in $\mathbb{R}^k$.

But we can do much better: using the exact formula for $A_4(n)$, we can prove the estimate

$$\sum_{n\leq x} A_4(n) = \rho_4 x^2 + O(x \log x).$$

Not only is this much more precise, but we can again use the formula for $A_4(n)$ to prove that the error term is as small as possible. Even better, this implies that, for all $k \geq 4$:

$$\sum_{n\leq x} A_k(n) = \rho_k x^{\frac{k-1}{2}} + O(x^{\frac{k}{2}-1} \log x),$$

which is, again, optimal. For proofs of these facts, see [3, Chapter 1.5].

As for the cases $k = 2, 3$, we can still do better than the geometric estimate we mentioned. Namely, we have:

$$\sum_{n\leq x} A_2(n) = \pi x + O(x^{\frac{1}{3}}),$$

and

$$\sum_{n\leq x} A_3(n) = \frac{4\pi}{3} x^{\frac{3}{2}} + O(x^{\frac{3}{4}}),$$

(see [3, Corollary 4.9]). However, the problem of obtaining the best possible estimates for $k = 2$ and $k = 3$ is still open.

Another concrete example is the following. $A_4(n)$ is the number of integral quaternions of squared norm n. Using this fact, and the correspondence between unit quaternions and rotations in $\mathbb{R}^3$, we can construct explicit free subgroups of $SO(3)$ of rank $\frac{p+1}{2}$, for a prime $p \equiv 1 \mod 4$. This is the main ingredient to prove the Banach-Tarski paradox, about the existence of paradoxical decompositions for the action of $SO(3)$ on S^2 [4, Chapter 2].

However, there are easier ways to find explicit free subgroups of $SO(3)$, so this would not be a sufficient motivation by itself. The real interest of these specific subgroups is that they were exploited by Lubotzky, Phillips and Sarnak to provide the first explicit constructions of Ramanujan graphs, in 1986 and 1988. These are expanding graphs (infinite families of finite graphs that are very well connected, despite having relatively few edges) that are optimal in a precise sense. This construction was a huge breakthrough in the theory of expanding graphs, that is of prime interest for both computer scientists and group theorists. Indeed, on the one hand, expanding graphs are very useful to build efficient networks; on the other hand, they all arise as Cayley graphs of finite quotients of groups with an interesting dynamical property (called property (τ), a generalization of Kazhdan's property (T)). See [4] for a thorough treatment (in particular chapter 7 for the construction of LPS graphs).

1.1 Idea of the proof

We will follow closely the approach taken in [1, Chapter VII]. To characterize $A_k(n)$, the idea is to encode them in a generating function

$$f(z) = \sum_{n=0}^{\infty} A_k(n) e^{\pi i n z},$$

and identify this function with another one, whose Fourier series we know well. For this we will use what we have already seen in talk 4.

Definition 1.2. The Jacobi theta function is defined by

$$\theta(z) := \sum_{n=-\infty}^{\infty} e^{\pi i n^2 z} = \sum_{n=-\infty}^{\infty} h^{n^2},$$

where $h = e^{\pi i z}$. Note that our definition differs from the original one by a imaginary part instead of a minus sign, i.e., the original function is our $\theta(iz)$. This will not impact the properties that we will use in the talk, except that this function is now defined and holomorphic on $\mathbb{H}$ instead of $\{z \in \mathbb{C} : Re(z) > 0\}$.

Using usual multiplication of series, one immediately computes that

$$\theta(z)^k = \left(\sum_{n=-\infty}^{\infty} h^{n^2} \right)^k = \sum_{n=0}^{\infty} A_k(n) h^n.$$

Therefore, our goal will be to find formulas for θ^4 and θ^8. For this, we will characterize the function θ by some of its properties and use Eisenstein series to create functions satisfying theses characterizations.

2 Sums of eight squares

2.1 A characterization of the theta function

We first recall some properties of θ and prove a few other ones. In what follows, we will always use the notation $z = x + iy$.

Proposition 2.1. *The function $\theta : \mathbb{H} \to \mathbb{C}$ satisfies the following formulas:*

(a) $\theta(z+2) = \theta(z)$*;*

(b) $\theta(-1/z) = \sqrt{z/i}\theta(z)$*;*

(c) $\lim_{y\to\infty} \theta(z) = 1$*;*

(d) $\lim_{y\to\infty} \sqrt{\frac{i}{z}}\theta(1-\frac{1}{z})e^{-\frac{\pi i z}{4}} = 2$*;*

(e) θ *does not vanish on* $\mathbb{H}$*.*

Proof. (a) Obvious since $e^{2\pi i n^2} = 1$ for any integer n.

(b) This is (almost) a reformulation of Proposition 3.3 of talk 4. The proof can also be done in the same way as point (d).

(c) This is a reformulation of Corollary 3.11 of talk 4.

(d) To prove this, we will use the formula

$$\theta\left(1-\frac{1}{z}\right) = \sqrt{\frac{z}{i}}\sum_{n=-\infty}^{\infty} e^{\pi i z(n+\frac{1}{2})^2}.$$

We follow the same idea as in Proposition 3.3 of talk 4. We denote, for a fixed z, the functions $f(x) := e^{-\pi x^2}$ and $g(x) := e^{\pi i z x^2} = f(\sqrt{-iz}x)$. Usual properties of the Fourier transform give us that

$$\hat{g}(x) = \frac{1}{\sqrt{-iz}}\hat{f}\left(\frac{x}{\sqrt{-iz}}\right).$$

We use the following version of Poisson summation formula which can be proven in the same way as the usual one:

$$\sum_{n=-\infty}^{\infty} f(n+y) = \sum_{n=-\infty}^{\infty} \hat{f}(n)e^{-2\pi i n y}.$$

We take $y = \frac{1}{2}$, so $g(n+y) = e^{\pi i z(x+\frac{1}{2})^2}$. Since $\hat{f} = f$, using the Poisson summation formula, we get

$$\sum_{n=-\infty}^{\infty} e^{\pi i z(n+\frac{1}{2})^2} = \frac{1}{\sqrt{-iz}}\sum_{n=-\infty}^{\infty} e^{-\pi\left(\frac{n}{\sqrt{-iz}}\right)^2} e^{-\pi i n} = \sqrt{\frac{i}{z}}\sum_{n=-\infty}^{\infty} e^{-\pi i\frac{n^2}{z}}e^{-\pi i n}.$$

We just have to notice that $e^{-\pi i n} = e^{\pi i n^2}$ for any $n \in \mathbb{Z}$. We conclude that

$$\sum_{n=-\infty}^{\infty} e^{\pi i z(n+\frac{1}{2})^2} = \sqrt{\frac{i}{z}}\sum_{n=-\infty}^{\infty} e^{\pi i n^2\left(1-\frac{1}{z}\right)} = \sqrt{\frac{i}{z}}\theta\left(1-\frac{1}{z}\right).$$

Now, from the formula, we can deduce that

$$\lim_{y\to\infty} \sqrt{\frac{z}{i}}^{-1} \theta(1-\frac{1}{z})e^{-\frac{\pi i z}{4}} = \lim_{y\to\infty} \sum_{n=-\infty}^{\infty} e^{\pi i z(n^2+n)}.$$

Then the result follows, since the sum converges uniformly (for $y \geq C > 0$, see the proof of Proposition 3.2 of talk 4) and each term except for $n = -1$ and $n = 0$ goes to 0.

(e) We will see later that we only need to prove that $\theta(z)$, $\theta(z+1)$ and $\theta(1-1/z)$ do not vanish in the usual fundamental domain $\mathcal{F}$ of $SL(2,\mathbb{Z})$. Since the minimal imaginary part in $\mathcal{F}$ is $\frac{\sqrt{3}}{2}$, we get for the first one

$$|\theta(z)-1| \leq 2\sum_{n=1}^{\infty} e^{-\pi n^2 y} \leq 2\sum_{n=1}^{\infty} e^{-\pi n\frac{\sqrt{3}}{2}} = 2\frac{e^{-\pi\frac{\sqrt{3}}{2}}}{1-e^{-\pi\frac{\sqrt{3}}{2}}} \leq 0.2$$

So $\theta(z)$ does not vanish in $\mathcal{F}$. The same computation works for the second one, since $|e^{\pi i n^2}| = 1$. For the last one, we use again the formula

$$\theta\left(1-\frac{1}{z}\right) = \sqrt{\frac{z}{i}} \sum_{n=-\infty}^{\infty} e^{\pi i z(n+\frac{1}{2})^2}.$$

So, up to a factor that does not vanish in $\mathbb{H}$, we get

$$\sum_{n=-\infty}^{\infty} e^{\pi i z(n^2+n)}$$

and one can use the same argument as before.

$\square$

What will see now is that in fact these four first properties characterize (powers of) the theta function in some sense.

Proposition 2.2. *Let $r \in \mathbb{Z}$, and let $f : \mathbb{H} \to \mathbb{C}$ be a holomorphic function with the following properties:*

(a) $f(z+2) = f(z)$;

(b) $f(-\frac{1}{z}) = \sqrt{\frac{z}{i}}^r f(z)$;

(c) $C := \lim_{y\to\infty} f(z)$ *exists;*

(d) $\lim_{y\to\infty} \sqrt{\frac{z}{i}}^{-r} f(1-\frac{1}{z})e^{-\frac{\pi i r z}{4}}$ *exists.*

Then $f(z) = C \cdot \theta^r(z)$.

Remark. We will prove that $f(z) = C \cdot \theta^r(z)$ for *some* constant C. Taking limits on both sides, it follows that $C = \lim_{y\to\infty} f(z)$.

Proof. To prove this proposition, we first reduce to the case $r = 0$. We set $\tilde{f}(z) := \frac{f(z)}{\theta^r(z)}$. Since θ doesn't vanish anywhere, we immediately see that $\tilde{f}$ is holomorphic. Moreover, it always satisfies the previous assumptions for $r = 0$. Hence, we just need to prove that $\tilde{f}$ is a constant function and the proposition follows. This is what the next proposition tell us. □

Proposition 2.3. *Let $f : \mathbb{H} \to \mathbb{C}$ be an holomorphic function with the properties:*

(a) $f(z+2) = f(z)$*;*

(b) $f(-\frac{1}{z}) = f(z)$*;*

(c) Both limits $a := \lim_{y\to\infty} f(z)$ *and* $b := \lim_{y\to\infty} f(1-1/z)$ *exist.*

Then f is a constant function.

To prove this proposition, we first need to introduce a subgroup of $SL(2,\mathbb{Z})$.

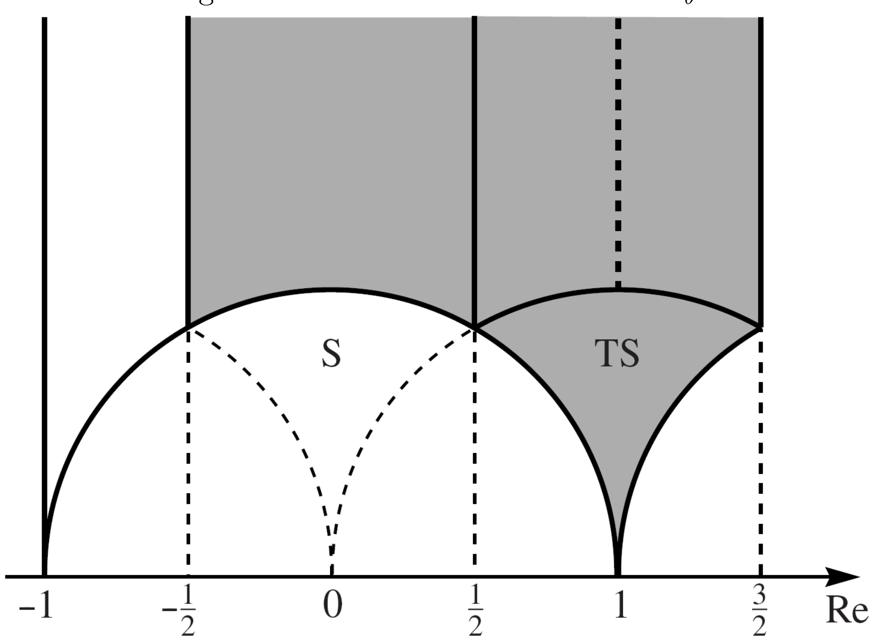
Definition 2.4. The theta group Γ_θ is the subgroup of $SL(2,\mathbb{Z})$ generated by the two matrices:

$$\Gamma_\theta := \left\langle \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \right\rangle = \left\langle T^2, S \right\rangle .$$

It is also characterized as the group of all matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2,\mathbb{Z}))$ satisfying

$$a+b+c+d = 0 \mod 2.$$

Figure 1: The fundamental domain $\mathcal{F}_\theta$.



Lemma 2.5.

$$\mathcal{F}_\theta := \mathcal{F} \cup \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \mathcal{F} \cup \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \mathcal{F}$$

is a fundamental domain for the theta group Γ_θ, where $\mathcal{F}$ is the usual fundamental domain for $SL(2,\mathbb{Z})$.

Remark. 1. From this, we can deduce what was missing in Proposition 2.1 (e) for the non-vanishing of the function θ : if $z \in \mathcal{F}$, then z, $z+1$ and $1-\frac{1}{z}$ cover the three parts of $\mathcal{F}_\theta$ and since θ is invariant under Γ_θ, we cover any $z \in \mathbb{H}$.

2. This fundamental domain is given by the equations

$$\mathcal{F}_\theta = \left\{ z \in \mathbb{H} \middle| -\frac{1}{2} \leq Re(x) \leq \frac{3}{2},\ 1 \leq |z| \text{ and } 1 \leq \left| z - \frac{3}{2} \right| \right\}.$$

Proof. We will only need that $\Gamma_\theta \cdot \mathcal{F}_\theta = \mathbb{H}$ so we will only prove that for any $z \in \mathbb{H}$, there exists some $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_\theta$ such that $Az \in \mathcal{F}_\theta$. For this, recall that $Im(Az) = \frac{Im(z)}{|cz+d|^2}$. Varying A, we get a discrete set of points given by all the possible values of $cz+d$. Between all of them, there musts exist one of minimal modulus, given by some matrix A. Multiplying by a multiple of T^2, we can make sure that $|Re(z)| \leq 1$. The way we chose A tells us in particular that

$$Im(Az) \geq Im(SAz) = \frac{Im(Az)}{|Az|^2},$$

i.e., $|Az| \geq 1$. We have proved that for any $z \in \mathbb{H}$, there exists a matrix $A \in \Gamma_\theta$ such that Az is in the region given by $-1 \leq Re(z) \leq 1$ and $|z| \geq 1$. In particular, $z \in \mathcal{F}_\theta$ if $-\frac{1}{2} \leq Re(z)$. In the other case, we just have to shift Az by T^2 to get the result. □

Proof of 2.3. We define

$$H(z) := (f(z)-a)(f(z)-b).$$

By assumption, we have that

$$\lim_{y\to\infty} H(z) = 0 \text{ and } \lim_{\mathcal{F}_\theta \ni z \to 1} H(z) = 0.$$

Since H is continuous, it must be bounded in an open neighbourhood of infinity and also in an open neighbourhood of 1 inside $\mathcal{F}_\theta$. If we remove these two neighbourhood, what is left is a compact subset of Γ_θ where, by continuity, H must be also bounded. But we also have $H(z+2) = H(z)$ and $H(-1/z) = H(z)$ so H is invariant the action Γ_θ hence it is bounded in $\mathbb{H}$. Therefore, by the maximum principle, it must be constant. Looking at the limits before, we see that H is zero everywhere. So h take values a or b. Since it is continuous on $\mathbb{H}$, it must be constant. □

2.2 Proof of the formula

In this section, we want to apply Proposition 2.2 to find a formula for θ^8 and hence for $A_8(n)$. For this, we need a function satisfying the following:

(a) $f(z+2) = f(z)$;

(b) $f(-\frac{1}{z}) = z^4 f(z)$;

(c) $\lim_{y\to\infty} f(z)$ exists;

(d) $\lim_{y\to\infty} z^{-4} f(1-\frac{1}{z}) e^{-2\pi i z}$ exists.

A first candidate would be $f = G_4$. We already saw in the proof of Theorem 2.4 of talk 2 that (a) and (b) hold since G_4 is a modular form and moreover that $\lim_{y\to\infty} f(z) = 2\zeta(4)$. Unfortunately, $G_4(1-\frac{1}{z}) = G_4(-\frac{1}{z}) = z^4 G_4(z)$ and so

$$z^{-4}G_4(1-\frac{1}{z})e^{-2\pi i z} = G_4(z)e^{-2\pi i z}.$$

But we just saw that G_4 has a non-zero limit as $y \to \infty$ and $|e^{-2\pi i z}| = e^{2\pi y}$ so the limit doesn't exist. A way to rectify this argument is to consider the following function:

Proposition 2.6. *For any integer $k > 2$, the function $g_k(z) := G_k(\frac{z+1}{2})$ satisfies the following:*

(a) $g_k(z+2) = g_k(z)$;

(b) $g_k(-1/z) = z^k g_k(z)$;

(c) $\lim_{y\to\infty} g_k(z) = 2\zeta(k)$.

Proof. The first and last formulas are again clear from what we already saw on G_k in talk 2. For the second formula, consider

$$A := \begin{pmatrix} 1 & -1 \\ 2 & -1 \end{pmatrix} \in SL(2,\mathbb{Z}).$$

Then

$$A \cdot \frac{z+1}{2} = \frac{\frac{z+1}{2}-1}{2\frac{z+1}{2}-1} = \frac{\frac{z-1}{2}}{z} = \frac{1}{2} - \frac{1}{2z} = \frac{-\frac{1}{z}+1}{2}.$$

Hence, by modularity of G_k, we conclude

$$g_k(-1/z) = G_k\left(A \cdot \frac{z+1}{2}\right) = \left(2\frac{z+1}{2} - 1\right)^k G_k(\frac{z+1}{2}) = z^k g_k(z).$$

□

Again, g_4 does not satisfy the last condition since

$$g_4(1-1/z) = G_4(-1/2z) = (2z)^4 G_4(2z),$$

so the same argument will work as for G_4.

To get out of this convergence problem, the idea is to take a linear combination of G_4 and g_4 in such a way that the limit of the combination exists. Let $a, b \in \mathbb{C}$ and

$$f(z) := aG_4(z) + bg_4(z).$$

The three first properties obviously hold by linearity and $\lim_{y\to\infty} f(z) = 2(a+b)\zeta(4)$. We also have already computed

$$z^{-4}f(1-1/z)e^{-2\pi i z} = (aG_4(z) + 16bG_4(2z))e^{-2\pi i z}.$$

Denote $q = e^{2\pi i z}$. We saw in talk 2 that G_4 admits a Fourier series, i.e., there exist coefficients $a_0, a_1, a_2, \ldots$ such that

$$G_4(z) = a_0 + a_1 q + a_2 q^2 + \ldots$$

Rewriting the equality above, we get

$$z^{-4}f(1-1/z)e^{-2\pi iz}=q^{-1}(a_0(a+16b)+O(q)).$$

Hence, if $a+16b=0$, the series converges as $y\to\infty$ since then $q\to 0$. We conclude that there exists a non-zero constant C such that

$$f(z)=C\theta^8(z).$$

Explicitly, $C=2(a+b)\zeta(4)$. We would like to have $C=1$. Combining with $a+16b=0$, we get that $b=-\frac{1}{30\zeta(4)}$ and $a=\frac{8}{15\zeta(4)}$. We proved the identity (recall that $\zeta(4)=\frac{\pi^4}{90}$):

$$\theta^8(z)=\frac{16}{30\zeta(4)}G_4(z)-\frac{1}{30\zeta(4)}G_4\left(\frac{z+1}{2}\right)=\frac{3}{\pi^4}\left(16G_4(z)-G_4\left(\frac{z+1}{2}\right)\right).$$

Theorem 2.7 (C.G.J. Jacobi, 1829)**.** *For any $n\in\mathbb{N}$, we have*

$$A_8(n)=16\sum_{d|n}(-1)^{n-d}d^3.$$

Proof. We just have to identify the Fourier series of G_4 and θ^8. We saw in talk 2 that

$$G_4(z)=2\zeta(4)+\frac{16\pi^4}{3}\sum_{n=1}^{\infty}\sigma_3(n)e^{2\pi inz}.$$

Hence, the identity above can be rewritten as

$$\sum_{n=0}^{\infty}A_8(n)h^n=1+16^2\sum_{n=1}^{\infty}\sigma_3(n)h^{2n}-16\sum_{n=1}^{\infty}\sigma_3(n)h^n(-1)^n.$$

where $h:=e^{\pi iz}$. Hence, we immediately get $A_8(0)=1$ and for the odd n, $A_8(n)=16\sigma_3(n)$ which is what we want since $n-d$ will always be even for any $d|n$. For the even n, we write $n=2^rm$, where m is odd, to get

$$A_8(n)=16^2\sigma_3(2^{r-1}m)-16\sigma_3(2^rm)=16^2\sum_{d|2^{r-1}m}d^3-16\sum_{d|2^rm}d^3.$$

To conclude, we group all the powers of two of the same divisor of m.

$$A_8(n)=16^2\sum_{d|m}(d^3+(2d)^3+...+(2^{r-1}d)^3)-16\sum_{d|m}(d^3+(2d)^3+...+(2^rd)^3)$$

$$=2\cdot 16\sum_{d|m}((2d)^3+(4d)^3+...+(2^rd)^3)-16\sum_{d|m}(d^3+(2d)^3+...+(2^rd)^3)$$

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

Because the only divisors that get a minus sign are the odd ones. $\square$

Remark. Actually, using that $a=\frac{8}{15\zeta(4)}$, $b=-\frac{1}{30\zeta(4)}$ and comparing the first h-coefficient, this expression for θ^8 provides a proof that $\zeta(4)=\frac{\pi^4}{90}$.

3 Sums of four squares

The goal of the second part of the talk is to prove the following formula for $A_4(n)$:

Theorem 3.1 (C.G.J. Jacobi, 24.4.1828)**.**

$$A_4(n) = 8 \sum_{4 \nmid d | n} d.$$

Looking at the proof of the formula for $A_8(n)$, the most natural thing to do would be to try and find an expression for θ^4 using the Eisenstein series of weight 2. However, a little more work has to be done in order to prove the properties in Proposition 2.2, since G_2 does not converge absolutely, so we need to be more careful when looking at the effect of fractional linear transformations on it. The first subsection will be devoted to developing the necessary theory to establish those properties. Then we will be able to proceed in an analogous way to the previous section.

3.1 The Eisenstein series of weight 2

We start by recalling what we already know from talk 2. We can define the Eisenstein series of weight 2 as:

$$G_2(z) := \sum_{m=-\infty}^{\infty} \left\{ \sum_{\substack{n=-\infty \\ n \neq 0 \text{ if } m=0}}^{\infty} (mz+n)^{-2} \right\}.$$

Even though G_k is absolutely convergent for $k > 2$, so the order of summation does not matter, here the brackets must be taken into account. Forcing this order of summation, we have seen in talk 2 (definition 2.6) that

$$G_2(z) = 2\zeta(2) - 8\pi^2 \sum_{n=1}^{\infty} \sigma_1(n) q^n,$$

where $q = e^{2\pi i z}$. Furthermore, we saw that this sum converges absolutely to a holomorphic function. But we know from talk 3 (theorem 1.4) that there are no modular forms of weight 2, so G_2 cannot possibly be modular. Still, we can say something about the effect of applying a fractional linear transformation to G_2. The following result was already mentioned in talk 2 (lemma 2.7):

Proposition 3.2. *Let $z \in \mathbb{H}$ and $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2,\mathbb{Z})$. Then*

$$G_2\left(\frac{az+b}{cz+d}\right) = (cz+d)^2 G_2(z) - 2\pi i c(cz+d).$$

In particular:

$$G_2\left(-\frac{1}{z}\right) = z^2 G_2(z) - 2\pi i z$$

and

$$G_2(z+1) = G_2(z).$$

Proof. We follow [2, Proposition 6 p. 19].

We will exploit the fact that, even though the Eisenstein series of weight 2 does not converge absolutely, it is extremely close to doing that. We will therefore insert a convergence factor and let it go to 0 to deduce the theorem.

For any $\varepsilon \in \mathbb{R}$ (we will shortly use the behaviour also for negative ε), we define:

$$G_{2,\varepsilon}(z) := \sum_{(0,0)\neq(m,n)\in\mathbb{Z}^2} \frac{1}{(mz+n)^2|mz+n|^{2\varepsilon}}.$$

The argument used in talk 2 to show absolute convergence and modularity of $G_k(z)$, for $k > 2$, can be applied to show that for all $\varepsilon > 0$, this sum converges absolutely and satisfies:

$$G_{2,\varepsilon}\left(\frac{az+b}{cz+d}\right) = (cz+d)^2|cz+d|^{2\varepsilon}G_{2,\varepsilon}(z).$$

Therefore, if $G_2^*(z) := \lim_{\varepsilon\to 0} G_{2,\varepsilon}(z)$ exists, it will also transform like a weight 2 modular function. The heart of the proof is showing that the limit exists, and can be expressed in terms of $G_2(z)$.

To this end, we introduce the following auxiliary function:

$$I_\varepsilon(z) = \int_{-\infty}^{\infty} \frac{dt}{(z+t)^2|z+t|^{2\varepsilon}} = \sum_{n=-\infty}^{\infty} \int_n^{n+1} \frac{dt}{(z+t)^2|z+t|^{2\varepsilon}}.$$

Notice that this integral is well-defined for $z \in \mathbb{H}$, and converges absolutely for $z \in \mathbb{H}$ and $\varepsilon > -1/2$. We introduce this function because the following difference is particularly well-behaved:

$$G_{2,\varepsilon}(z) - 2\sum_{m=1}^{\infty} I_\varepsilon(mz) = 2\sum_{n=1}^{\infty}\frac{1}{n^{2+2\varepsilon}} + 2\sum_{m=1}^{\infty}\sum_{n=-\infty}^{\infty}\frac{1}{(mz+n)^2|mz+n|^{2\varepsilon}} - 2\sum_{m=1}^{\infty} I_\varepsilon(mz).$$

We know that the sums defining $G_{2,\varepsilon}$ are absolutely convergent, for ε. The last sum appearing is also absolutely convergent, for $\varepsilon > 0$. Indeed, setting $u = t/m$, we have:

$$I_\varepsilon(mz) = \int_{-\infty}^{\infty} \frac{dt}{(mz+t)^2|mz+t|^{2\varepsilon}} = \int_{-\infty}^{\infty} \frac{m\,du}{(mz+mu)^2|mz+mu|^{2\varepsilon}} = \frac{1}{m^{1+2\varepsilon}} I_\varepsilon(z);$$

so

$$\sum_{m=1}^{\infty} I_\varepsilon(mz) = I_\varepsilon(z)\sum_{m=1}^{\infty}\frac{1}{m^{1+2\varepsilon}} = I_\varepsilon(z)\zeta(1+2\varepsilon).$$

Therefore for $\varepsilon > 0$, by absolute convergence of both sums we can write:

$$G_{2,\varepsilon}(z) - 2\sum_{m=1}^{\infty} I_\varepsilon(mz) = 2\sum_{n=1}^{\infty}\frac{1}{n^{2+2\varepsilon}} + 2\sum_{m=1}^{\infty}\sum_{n=-\infty}^{\infty}\left[\frac{1}{(mz+n)^2|mz+n|^{2\varepsilon}} - \int_n^{n+1}\frac{dt}{(mz+t)^2|mz+t|^{2\varepsilon}}\right].$$

The first sum converges absolutely for $\varepsilon > -1/2$. As for the second one, let $f(t)$ be the integrand. Recall that by the mean value theorem, for $n \le t \le (n+1)$, we have $|f(t) - f(n)| \le \sup_{n \le u \le (n+1)} |f'(u)| = O(|mz+n|^{-3-2\varepsilon})$. Therefore the absolute value of the summand can be estimated as follows:

$$|f(n) - \int_n^{n+1} f(t)dt| = |\int_n^{n+1} f(n) - f(t)dt| \le \int_n^{n+1} |f(t) - f(n)|dt = O(|mz+n|^{-3-2\varepsilon}).$$

This shows that the second sum also converges absolutely for $\varepsilon > -1/2$. This allows us to calculate the limit of the right-hand-side as $\varepsilon \to 0$ simply by plugging in $\varepsilon = 0$. This gives:

$$2\zeta(2) + 2\sum_{m=1}^{\infty} \sum_{n=-\infty}^{\infty} \left[\frac{1}{(mz+n)^2} - \int_n^{n+1} \frac{dt}{(mz+t)^2} \right].$$

Recall that, for a fixed m, the sum over $n \in \mathbb{Z}$ converges absolutely, so this allows us to change the order of summation and get:

$$2\zeta(2) + 2\sum_{m=1}^{\infty} \left[\sum_{n=-\infty}^{\infty} \frac{1}{(mz+n)^2} - \int_{-\infty}^{\infty} \frac{dt}{(mz+t)^2} \right] = 2\zeta(2) + 2\sum_{m=1}^{\infty} \sum_{n=-\infty}^{\infty} \frac{1}{(mz+n)^2} = G_2(z),$$

where we used that

$$\int_{-\infty}^{\infty} \frac{dt}{(mz+t)^2} = -\frac{1}{mz+t}\bigg|_{-\infty}^{\infty} = 0.$$

To recapitulate, so far we have:

$$G_2(z) = \lim_{\varepsilon \to 0} \left(G_{2,\varepsilon}(z) - 2\sum_{m=1}^{\infty} I_\varepsilon(mz) \right).$$

Now we need to estimate the sum to the right-hand-side of this equality. For $\varepsilon > -1/2$:

$$I_\varepsilon(x+iy) = \int_{-\infty}^{\infty} \frac{dt}{(x+t+iy)^2(|x+t+iy|^2)^\varepsilon} = \int_{-\infty}^{\infty} \frac{dt}{((x+t)+iy)^2((x+t)^2+y^2)^\varepsilon}.$$

We then set $u = (x+t)/y$ (recall that $y > 0$) to get:

$$I_\varepsilon(x+iy) = \int_{-\infty}^{\infty} \frac{y\,du}{(yu+iy)^2(y^2u^2+y^2)^\varepsilon} = \frac{1}{y^{1+2\varepsilon}} \int_{-\infty}^{\infty} \frac{du}{(u+i)^2(u^2+1)^\varepsilon} =: \frac{1}{y^{1+2\varepsilon}} I(\varepsilon).$$

Therefore, for all $\varepsilon > 0$:

$$\sum_{m=1}^{\infty} I_\varepsilon(mz) = \sum_{m=1}^{\infty} \frac{1}{(my)^{1+2\varepsilon}} I(\varepsilon) = \frac{1}{y^{1+2\varepsilon}} I(\varepsilon) \sum_{m=1}^{\infty} \frac{1}{m^{1+2\varepsilon}} = \frac{1}{y^{1+2\varepsilon}} I(\varepsilon)\zeta(1+2\varepsilon).$$

We are left to estimate this last term. First, we estimate $I(\varepsilon)$. We have:

$$I(0) = \int_{-\infty}^{\infty} \frac{du}{(u+i)^2} = -\frac{1}{u+i}\Big|_{-\infty}^{\infty} = 0.$$

Using the Leibniz integral rule to differentiate an integral function, we also have:

$$I'(0) = \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \int_{-\infty}^{\infty} \frac{du}{(u+i)^2(u^2+1)^\varepsilon} = \int_{-\infty}^{\infty} \left(\frac{\partial}{\partial\varepsilon}\Big|_{\varepsilon=0} \frac{1}{(u+i)^2(u^2+1)^\varepsilon} \right) du =$$

$$= \int_{-\infty}^{\infty} \frac{-\log(u^2+1)}{(u+i)^2} du = \left(\frac{1+\log(u^2+1)}{u+i} - \arctan(u) \right)\Big|_{-\infty}^{\infty} = -\pi.$$

So $I(\varepsilon) = -\pi\varepsilon + O(\varepsilon^2)$.

As for the second factor, we know from talk 4 (theorem 4.1) that ζ has a simple pole at 1 with residue 1, so its Laurent expansion around 1 is $\zeta(s) = \frac{1}{s-1} + O(1)$. Applying this to $s = (1+2\varepsilon)$ we obtain $\zeta(1+2\varepsilon) = \frac{1}{2\varepsilon} + O(1)$. Therefore:

$$\sum_{m=1}^{\infty} I_\varepsilon(mz) = \frac{1}{y^{1+2\varepsilon}} I(\varepsilon)\zeta(1+2\varepsilon) = \frac{1}{y^{1+2\varepsilon}} \left(-\pi\varepsilon + O(\varepsilon^2)\right) \left(\frac{1}{2\varepsilon} + O(1) \right) =$$

$$= \frac{1}{y^{1+2\varepsilon}} \frac{-\pi}{2} \left(1 + O(\varepsilon)\right) \xrightarrow[\varepsilon\to 0]{} -\frac{\pi}{2y}.$$

We conclude that

$$G_2^*(z) = \lim_{\varepsilon\to 0} G_{2,\varepsilon}(z) = G_2(z) + \lim_{\varepsilon\to o} 2\sum_{m=1}^{\infty} I_\varepsilon(mz) = G_2(z) - \frac{\pi}{y}.$$

Moreover, since this limit exists, it transforms like a weight 2 modular form. Therefore we obtain the final result with one last calculation:

$$G_2\left(\frac{az+b}{cz+d}\right) = G_2^*\left(\frac{az+b}{cz+d}\right) + \frac{\pi}{Im\left(\frac{az+b}{cz+d}\right)} = (cz+d)^2 G_2^*(z) + \frac{\pi|cz+d|^2}{y} =$$

$$= (cz+d)^2 G_2(z) + \frac{\pi}{y}\left(|cz+d|^2 - (cz+d)^2\right) = (cz+d)^2 G_2(z) + \frac{\pi(cz+d)}{y}\left((c\bar{z}+d) - (cz+d)\right) =$$

$$= (cz+d)^2 G_2(z) + \frac{\pi(cz+d)}{y}(2ciy) = (cz+d)^2 G_2(z) - 2\pi i c(cz+d).$$

□

3.2 Proof of the formula

Analogously to the previous section, we want to use Proposition 2.2 to find an expression for θ^4. This time we look at:

$$f(z) := aG_2\left(\frac{z}{2}\right) + bG_2(2z).$$

We need to choose the coefficients a and b so that the following are satisfied:

(a) $f(z+2) = f(z)$;

(b) $f(-\frac{1}{z}) = -z^2 f(z)$;

(c) $\lim\limits_{y\to\infty} f(z)$ exists;

(d) $\lim\limits_{y\to\infty} z^{-2} f(1-\frac{1}{z})e^{-\pi i z}$ exists.

Moreover, in order to obtain θ^4, we want the limit in (c) to be equal to 1.

(a) is satisfied, since G_2 is already 1-periodic (proposition 3.2). To obtain (b), we plug in the formula from Proposition 3.2:

$$f\left(-\frac{1}{z}\right) = aG_2\left(-\frac{1}{2z}\right) + bG_2\left(-\frac{1}{z/2}\right) = a\left(4z^2G_2(2z) - 4\pi i z\right) + b\left(\frac{z^2}{4}G_2\left(\frac{z}{2}\right) - \pi i z\right) =$$

$$= z^2\left(4aG_2(2z) + \frac{b}{4}G_2\left(\frac{z}{2}\right)\right) - \pi i z(4a+b).$$

We want to get rid of the error term, which forces us to set $b = -4a$. This actually works out for the z^2 term as well, and we obtain:

$$f(z) = a\left(G_2\left(\frac{z}{2}\right) - 4G_2(2z)\right);$$

$$f\left(-\frac{1}{z}\right) = -z^2 f(z).$$

To prove (c), we will use the Fourier expansion of G_2, which we recall for convenience:

$$G_2(z) = 2\zeta(2) - 8\pi^2\sum_{n=1}^{\infty}\sigma_1(n)q^n.$$

Remember from talk 2 that this converges absolutely. Therefore we can exchange limit and sum to get:

$$\lim_{y\to\infty}\sum_{n=1}^{\infty}\sigma_1(n)q^n = \sum_{n=1}^{\infty}\lim_{y\to\infty}\sigma_1(n)q^n = 0,$$

and so

$$\lim_{y\to\infty} G_2(z) = 2\zeta(2).$$

Applying this to f yields:

$$\lim_{y\to\infty} f(z) = a \lim_{y\to\infty} \left(G_2\left(\frac{z}{2}\right) - 4G_2(2z)\right) = -6a\zeta(2).$$

Therefore the limit exists, and setting $a = -\frac{1}{6\zeta(2)}$ we obtain (c), as we wanted.

It remains to prove condition (d). In order to calculate $f(1-\frac{1}{z})$, we need to look at $G_2(2-\frac{2}{z})$ and $G_2(\frac{1-1/z}{2})$. In both cases, we can use Proposition 3.2 to transform these expressions. For the first one, using periodicity and the formula for circle inversion, we have:

$$G_2\left(2-\frac{2}{z}\right) = G_2\left(-\frac{1}{z/2}\right) = \left(\frac{z}{2}\right)^2 G_2\left(\frac{z}{2}\right) - \pi i z.$$

As for the second one, we would like to use the transformation $z \mapsto \frac{1-1/z}{2} = \frac{z-1}{2z}$, but the determinant is 2, and not 1. So instead we calculate:

$$G_2\left(\frac{z-1}{2z-1}\right) = (2z-1)^2 G_2(z) - 4\pi i(2z-1).$$

Then, replacing z by $(z+1)/2$ we get:

$$G_2\left(\frac{1-1/z}{2}\right) = G_2\left(\frac{z-1}{2z}\right) = G_2\left(\frac{(z+1)/2-1}{2(z+1)/2-1}\right) = z^2 G_2\left(\frac{z+1}{2}\right) - 4\pi i z.$$

This gives us an expression for the limit we want to calculate:

$$z^{-2} f\left(1-\frac{1}{z}\right) = z^{-2} a \left(G_2\left(\frac{1-1/z}{2}\right) - 4G_2\left(2-\frac{2}{z}\right)\right) = a\left(G_2\left(\frac{z+1}{2}\right) - G_2\left(\frac{z}{2}\right)\right).$$

Now recall that $G_2(z)$ can be written as a power series in $q = e^{2\pi i z}$. Therefore, writing $h = e^{\pi i z}$, we have:

$$G_2\left(\frac{z}{2}\right) = 2\zeta(2) - 8\pi^2 \sum_{n=1}^{\infty} \sigma_1(n) h^n;$$

$$G_2\left(\frac{z+1}{2}\right) = 2\zeta(2) - 8\pi^2 \sum_{n=1}^{\infty} \sigma_1(n) h^n e^{\pi i} = 2\zeta(2) + 8\pi^2 \sum_{n=1}^{\infty} \sigma_1(n) h^n.$$

Combining these two (absolutely converging) power series with the previous result, we have:

$$z^{-2} f\left(1-\frac{1}{z}\right) h^{-1} = 16a\pi^2 \sum_{n=1}^{\infty} \sigma_1(n) h^{n-1} = 16a\pi^2 + 16a\pi^2 \sum_{n=1}^{\infty} \sigma_1(n+1) h^n.$$

Once again, by absolute convergence, the sum on the right goes to 0 as $Im(z) \to \infty$, so:

$$\lim_{y\to\infty} z^{-2} f\left(1-\frac{1}{z}\right) h^{-1} = 16a\pi^2$$

exists, which proves (d).

We have thus verified all the conditions necessary to apply Proposition 2.2. Recalling that $a = -\frac{1}{6\zeta(2)} = -\frac{1}{\pi^2}$, we conclude:

$$\theta^4(z) = f(z) = a\left(G_2\left(\frac{z}{2}\right) - 4G_2(2z)\right) = \frac{4G_2(2z) - G_2(z/2)}{\pi^2}.$$

Passing to the h-expansions, where as usual $h = e^{\pi i z}$:

$$\sum_{n=0}^{\infty} A_4(n)h^n = \frac{1}{\pi^2}\left(4\left(2\zeta(2) - 8\pi^2\sum_{n=1}^{\infty}\sigma_1(n)h^{4n}\right) - \left(2\zeta(2) - 8\pi^2\sum_{n=1}^{\infty}\sigma_1(n)h^n\right)\right) =$$

$$= 1 + 8\left(\sum_{n=1}^{\infty}\sigma_1(n)h^n - 4\sum_{\substack{n=1\\4|n}}^{\infty}\sigma_1\left(\frac{n}{4}\right)h^n\right) =: 1 + 8\sum_{n=1}^{\infty}a_n h^n.$$

Now let us look closely at second sum in the next-to-last expression. If $4|n$, then:

$$\sigma_1\left(\frac{n}{4}\right) = \sum_{d|\frac{n}{4}} d.$$

by definition. But instead of summing over the divisors of $\frac{n}{4}$, we can sum over the divisors of n that are also divisible by 4:

$$\sigma_1\left(\frac{n}{4}\right) = \sum_{4|d|n}\frac{d}{4}.$$

The advantage of this second expression is that it is also well-defined for n not divisible by 4, in which case it is zero, so we do not need to separate cases. Finally, comparing the coefficients, for all $n \geq 1$:

$$A_4(n) = 8a_n = 8\left(\sum_{d|n} d - 4\sum_{4|d|n}\frac{d}{4}\right) = 8\sum_{4\nmid d|n} d;$$

which is what we wanted to prove.

Remark. Using that $a = \frac{1}{6\zeta(2)}$ and comparing the first h-coefficient, this expression for θ^4 provides a proof that $\zeta(2) = \frac{\pi^2}{6}$.

4 Other formulae

We end this talk by discussing the general problem of finding explicit formulas for $A_k(n)$. Grosswald's book [5] provides a thorough treatment of what was known in the 80's, and more recently further progress has been made.

The problem of determining the number of representations of a number as a sum of squares is very old. Before modular forms were even introduced, the following theorems were proven by elementary methods.

Theorem 4.1 (Sums of two squares - Gauss, 1801)**.** *For an integer n, denote by $d(n)$ the number of divisors of n. Let $n = 2^f n_1 n_3$, where $n_1 = \prod_{p\equiv 1 \mod 4} p^{r_p}$ and $n_3 = \prod_{q\equiv 3 \mod 4} q^{s_q}$. Then:*

- *If some s_q is odd, then $A_2(n) = 0$.*
- *If all s_q are even, then $A_2(n) = 4d(n_1)$.*

For a proof, see [5, Chapter 2].

Theorem 4.2 (Sum of four squares - Lagrange, 1770)**.** *$A_4(n) \neq 0$ for all positive integers n.*

Of course this follows from Jacobi's four-square formula, but it was proven via elementary methods much earlier. For a proof, see [5, Chapter 3].

Theorem 4.3 (Sum of three squares - Legendre, 1798)**.** *$A_3(n) \neq 0$ if and only if n is not of the form $4^a b$, where $b \equiv 7 \mod 8$. Moreover, $A_3(4^a n) = A_3(n)$.*

As for the formula for the exact value of $A_3(n)$, it does exist but it is much more complicated. For proofs, see [5, Chapter 4].

To go beyond these results, some more sophisticated machinery has to be introduced. For instance, we proved formulas for $A_4(n)$ and $A_8(n)$ using modular forms. However, to get more formulas, a different method is needed: if we go on using the Eisenstein series, we run into trouble in higher degrees with the appearance of more modular forms, including cusp forms. It is still possible to use more complex versions of the theta function to prove formulas for $A_k(n)$, where $k \leq 12$ is even. For proofs see [5, Chapter 9].

There are other, related formulas that one can study. For example, one may wish to look at representations as sums of non-zero integer squares (see [5, Chapter 6]), or count essentially distinct representations, that is, representations that cannot be obtained from each other by permuting the summands or changing their sign (see [5, Chapter 7]).

The greatest breakthrough of the last years is a result of Milne [6], proven in 2000, giving exact formulas for $A_{4k^2}(n)$ and $A_{4k(k+1)}(n)$ for all k. The paper in which Milne proves these formulas is around 150 pages long, spanning a whole issue of the Ramanujan journal.

References

[1] E. Freitag, R. Busam: *Complex Analysis - Second Edition*. Springer, 2009.

[2] J.H. Bruinier, G. van der Geer, G. Harder, D. Zagier: *The 1-2-3 of Modular Forms - Lectures at a Summer School in Nordfjordeid, Norway* . Springer, 2008.

[3] H. Iwaniec, E. Kowalski: *Analytic Number Theory*. AMS, Colloquium Publications, Volume 53, 2004.

[4] A. Lubotzky: *Discrete Groups, Expanding Graphs and Invariant Measures - Second Edition*. Birkhäuser, 1994.

[5] E. Grosswald: *Representations of Integers as Sums of Squares*. Springer, 1985.

[6] S.C. Milne: *Infinite Families of Exact Sums of Squares Formulas, Jacobi Elliptic Functions, Continued Fractions, and Schur Functions*. The Ramanujan Journal, 6, 7–149, 2002.