

SUMS OF 6 SQUARES

ILA VARMA

We begin by recalling the general situation: For $n \in \mathbb{Z}_{>0}$ and $m \in \mathbb{Z}_{\geq 0}$, we define

$$r_{\mathbb{Z}^n}(m) = \{x \in \mathbb{Z}^n : x_1^2 + x_2^2 + \dots + x_n^2 = m\}.$$

We are interested in studying

$$\theta_{\mathbb{Z}^n} = \theta_n(q) = \sum_{m \geq 0} r_{\mathbb{Z}^n}(m) q^m.$$

From the first lecture, we know that in general θ_n have the following symmetries:

$$(\theta_1)^n = \theta_n \quad \theta_n(-1/z) = (iz)^{n/2} \theta_n(z) \quad \theta_n(z+2) = \theta_n(z).$$

The middle equality makes θ_n look like a modular form of weight $n/2$, but not of the full group $\mathrm{SL}_2(\mathbb{Z})$; instead we can see that it is “invariant” under the actions of the subgroup

$$\Gamma_1(4) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} \right\}.$$

Thus, these theta functions are of level 4. Let us remember the definition of a modular form on such a congruence subgroup.

Definition 1. A modular form of weight k with respect to Γ is a function $f : \mathbb{H} \rightarrow \mathbb{C}$ such that

- (1) f is holomorphic.
- (2) $f[\gamma]_k = f$ for all $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \gamma \in \Gamma$ where $f[\gamma]_k(\tau) := (c\tau + d)^{-k} f(\gamma(\tau))$.
- (3) $f[\alpha]_k$ is holomorphic at ∞ for all $\alpha \in \mathrm{SL}_2(\mathbb{Z})$.

We can think of holomorphic at ∞ in the following manner. There is a correspondence between functions on the upper-half plane and functions on the punctured disk:

$$\begin{aligned} f : \mathbb{H} &\longrightarrow \mathbb{C} & \xleftarrow{1-1} & g : D \setminus \{0\} \longrightarrow \mathbb{C} \\ f(\tau) &= & g(e^{2\pi i \tau / h}) \\ \text{holomorphic at } \infty &\longleftrightarrow & \text{extends to } 0 \end{aligned}$$

Here, h is the minimal integer such that $\begin{pmatrix} 1 & h \\ 0 & 1 \end{pmatrix} \in \Gamma$. The space of all such modular forms is denoted as $\mathcal{M}_k(\Gamma)$. It is a vector space over $\mathbb{C}$, and it can be shown that it is finite. In order to understand θ_n , we can study it as an element of $\mathcal{M}_{n/2}(\Gamma_1(4))$.

We are concerned with the case when $n = 6$. Then θ_6 is a weight 3 modular form on $\Gamma_1(4)$, i.e. an element of $\mathcal{M}_3(\Gamma_1(4))$. To understand this space, we first want to compute the dimension. Recall the modular curve associated to any Γ from the first few lectures. We take $\Gamma_1(4) \backslash (\mathbb{H} \cup \mathbb{P}_1(\mathbb{Q})) =: X_1(4)$. Note that it is isomorphic to the Riemann sphere, $\mathbb{P}_1(\mathbb{C})$. There is a geometric interpretation of modular forms as sections of line bundles on the moduli space, so we can consider ω , the line bundle of weight 1 forms on the space above. As for elliptic curves, we can calculate the degree of ω using the formula $\deg(\omega) = \frac{1}{24} [\mathrm{SL}_2(\mathbb{Z}) : \Gamma_1(4)]$. To calculate this index, note that $\Gamma(4)$, the set of all matrices in $\mathrm{SL}_1(\mathbb{Z})$ that are congruent (mod 4) to the identity matrix has index $4^3 (1 - \frac{1}{2^2}) = 48$. Consider the map $\Gamma_1(4) \rightarrow \mathbb{Z}/4\mathbb{Z}$ sending $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto b \pmod{4}$.

This is a surjective map with kernel $\Gamma(4)$, thus $[\Gamma_1(4) : \Gamma(4)] = 4$. This implies that $[\mathrm{SL}_2(\mathbb{Z}) : \Gamma_1(4)] = 12$, hence $\deg(\omega) = \frac{1}{2}$. (Another way to see this is to count the number of primitive elements in $\mathbb{Z}/4\mathbb{Z} \oplus \mathbb{Z}/4\mathbb{Z}$, which is 12.) Normally, since degrees are supposed to be integral, this technically implies that such a line

bundle doesn't exist, but it implies that 1 of the cusps is not "regular." (Note that there are three cusps for $\Gamma_1(4)$, i.e. three distinct conjugates of ∞ under $\Gamma_1(4)$ action.) Because we are interested in weight 3 forms, we want the degree of $\omega^{\otimes 3}$, which is easily calculated from here to be $\frac{3}{2}$. By Riemann-Roch, we can calculate

$$\dim(\mathcal{M}_3(\Gamma_1(4))) = \dim(H^0(X_1(4), (\omega)^{\otimes 3})) = \left\lfloor \frac{3}{2} \right\rfloor + 1 - 0 = 2.$$

In order to understand θ_6 as an element of this space, we want to write it in terms of a "good" basis for $\mathcal{M}_3(\Gamma_1(4))$. We try to find 2 independent series using a "natural" method.

Magical Fact. We can do so by connecting our question to the study of $\text{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$.

We have already seen how to characterize $\text{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ using abelian extensions of number fields via class field theory. We can go further to study $\text{Gal}(\mathbb{Q}/\mathbb{Q})$ by understanding its linear representations, i.e. linear maps

$$\text{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \longrightarrow \text{GL}_n(F)$$

where $F = \mathbb{C}, \mathbb{Q}_l$. In this language, we can regard the Artin map for any number field $K/\mathbb{Q}$ as a character $\chi : \text{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \rightarrow \text{GL}_1(\mathbb{C})$. Then $\ker(\chi)$ gives rise to the abelian extensions of $\mathbb{Q}$, and such 1-dimensional representations will factor through $\text{Gal}(F/\mathbb{Q}) \rightarrow \text{GL}_1(\mathbb{C})$ where F is this fixed field. However, this is not the end of the story; rather, we are more interested in larger dimensional representations as well as representations where $F \neq \mathbb{C}$. This analysis of $\text{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ via representation theory gives a way of understanding the structure of the group independent of your choice for $\bar{\mathbb{Q}}$.

Example. Recall the sum of 2 squares situation, where we introduced the Dirichlet character (of conductor 4)

$$(\mathbb{Z}/4\mathbb{Z})^\times \rightarrow \mathbb{C}^\times \quad \pm 1 \mapsto \pm 1.$$

We can view this as factoring a representation of $\text{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ through $\text{Gal}(\mathbb{Q}(i)/\mathbb{Q}) \cong (\mathbb{Z}/4\mathbb{Z})^\times$. We know that a prime p splits in $\mathbb{Q}(i)$ if and only if $p \equiv 1 \pmod{4}$, p is inert in $\mathbb{Q}(i)$ if and only if $p \equiv 3 \pmod{4}$, and $p = 2$ is the only prime that ramifies. Thus, $\chi : \text{Gal}(\mathbb{Q}(i)/\mathbb{Q}) \rightarrow \mathbb{C}^\times$ is this exact Dirichlet character described above when we identify $\text{Gal}(\mathbb{Q}(i)/\mathbb{Q})$ with $(\mathbb{Z}/4\mathbb{Z})^\times$. For $p = 2$, we can set $\chi(2) = 0$ and extend multiplicatively to all positive integers.

Note. Note that an odd integer can be written as a sum of 2 squares if and only if $\chi(m) = 1$. For the even case, we must first divide out by the higher power of 2 and then the same criterion holds.

This is a 1-dimensional representation of conductor 4. (We will see later that representations of conductor 4 correspond to modular forms on $\Gamma_1(4)$). If we want to consider 2-dimensional (reducible) representations of conductor 4, we can produce these by taking direct sums of 1-dimensional representations, e.g. $1 \oplus \chi$. For this representation, note that

$$\text{Tr}(\text{Frob}_p) = \text{Tr}(p) = \begin{cases} 2 & \text{if } p \equiv 1 \pmod{4} \\ 0 & \text{if } p \equiv 3 \pmod{4} \\ 1 & \text{if } p = 2 \text{ (no contribution from } \chi) \end{cases}$$

This is in fact the only 2-dimensional reducible representation of conductor 4 that we can construct as there are no Dirichlet characters of conductor 2 (exercise).

To every Dirichlet character, we can associate a generalization of the Riemann zeta function called an L -function defined as follows:

$$L(s, \chi) := \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s} = \prod_p \frac{1}{1 - \chi(p)p^{-s}}.$$

The product formula is known as the Euler product expansion. We do this because there is a nice relationship between L -functions of conductor 4 coming from such Galois representations and modular forms on $\Gamma_1(4)$

via the Mellin transform. Note that when $\chi = 1$, we get ζ back. For $\chi = \chi_4$ defined above,

$$\begin{aligned} L(s, \chi_4) &= 1 - \frac{1}{3^s} + \frac{1}{5^s} - \frac{1}{7^s} + \frac{1}{9^s} - \dots \\ &= \frac{1}{(1+3^{-s})} \frac{1}{(1-5^{-s})} \frac{1}{(1+7^{-s})} \dots \end{aligned}$$

For 2 dimensions, we use the fact that $L(s, \chi_1 \oplus \chi_2) = L(s, \chi_1)L(s, \chi_2)$. For the 2-dimensional representation $1 \oplus \chi_4$ given above, we have

$$L(s, 1 \oplus \chi_4) = \prod_p \frac{1}{1 - \chi(p)p^{-s}} \prod_p \frac{1}{1 - p^{-s}} = \sum_{n_1=1}^{\infty} \frac{\chi(n_1)}{n_1^s} \sum_{n_2=1}^{\infty} \frac{1}{n_2^s}.$$

Note that if we combine summations by summing over $n = n_1 n_2$ and count the number of times we sum over such n , we conclude that

$$L(s, 1 \oplus \chi_4) = \sum_{n \geq 1} \left(\sum_{d|n} \chi(d) \right) \frac{1}{n^s}$$

Via Mellin transform, this corresponds to $G(q) = c + \sum_{m \geq 1} \left(\sum_{d|m} \chi(d) \right) q^m$. This is a modular form of weight 1 (we will see why later), and we can see that computing the first few coefficients gives $\theta_2(q) = 4G(q)$ for the sum of 2 squares formula. Looking back at our dimension formula, it is easy to see that $\mathcal{M}_1(\Gamma_1(4))$ has dimension 1, hence G is our “good” basis and we have found a nice formula for the representation of integers by sum of 2 squares,

$$r_{\mathbb{Z}^2}(m) = 4 \sum_{d|m} \chi(d).$$

Weight 3. In the previous case, we have found a 2-dimensional representation of conductor 4 that somehow relates to weight 1 modular forms. To every Galois representation, we can associate a Hodge-Tate weight based on the characters in its decomposition and it turns out that 2-dimensional Galois representations of conductor 4 that arise as a direct sum of characters with Hodge-Tate weights $(0, k-1)$ correspond to weight k modular forms on $\Gamma_1(4)$. For the sum of 2 squares, we used $1 \oplus \chi$, both of which have Hodge-Tate weight 0. For the case of sum of 6 squares, i.e., $k = 3$, we must find a Galois representation with Hodge-Tate weights $(0, 2)$. To do so, we introduce the l -adic cyclotomic character: this is the unique character that factors through $\text{Gal}(\mathbb{Q}(\zeta_l^\infty)/\mathbb{Q})$ such that

$$\begin{aligned} \chi_l : \text{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) &\longrightarrow \mathbb{Z}_l^\times \hookrightarrow \text{GL}_1(\mathbb{Q}_l) \\ \sigma &\longmapsto \chi_l(\sigma) \quad \text{s.t. } \sigma(\zeta) = \zeta^{\chi(\sigma)} \quad \forall \zeta \in \mathbb{Q}(\zeta_l^\infty)_{\text{tor}}^\times \end{aligned}$$

This character is unramified at all primes $p \neq l$, hence it has Hodge-Tate weight 1, and furthermore for $\chi_l(\text{Frob}_p) = p$. Since we disregard l for l -adic representations, χ_l has conductor 1. Using χ_l as well as χ_4 and the trivial representation, we can build 2-dimensional (reducible) representations with conductor 4 and Hodge-Tate weight $(0, 2)$:

$$\chi_4 \oplus \chi_l^2 \quad 1 \oplus \chi_4 \chi_l^2.$$

Because there are no characters of conductor 2, these are the only two possibilities. Note that because these representations are independent, we will be able to produce two independent modular forms of weight 3 on $\Gamma_1(4)$.

The L -series associated to these two representations are as follows:

$$L_1(s, \chi_4 \oplus \chi_l^2) = \prod_p \frac{1}{1 - \chi_4(p)p^{-s}} \prod_p \frac{1}{1 - p^2 p^{-s}} \quad L_2(s, 1 \oplus \chi_4 \chi_l^2) = \prod_p \frac{1}{1 - p^{-s}} \prod_p \frac{1}{1 - \chi_4(p)p^2 p^{-s}}.$$

In summation formula,

$$\begin{aligned}
L_1(s) &= \sum_{n_1 \geq 1} \chi(n_1) n_1^{-s} \sum_{n_2 \geq 0} n_2^{2-s} = \sum_{n_1, n_2 \geq 1} \chi(n_1) n_2^2 (n_1 n_2)^s = \sum_{n \geq 1} \left(\sum_{dd'=n} \chi(d) d'^2 \right) n^{-s} \\
L_2(s) &= \sum_{n_1 \geq 1} n_1^{-s} \sum_{n_2 \geq 1} \chi_4(n_2) n_2^2 n_2^{-s} = \sum_{n_1, n_2 \geq 1} \chi(n_2) n_2^2 (n_1 n_2)^{-s} = \sum_{n \geq 1} \left(\sum_{d|n} \chi(d) d^2 \right) n^{-s}
\end{aligned}$$

Applying the Mellin transform as described in the first lecture, we get

$$G_1(q) = c_1 + \sum_{n \geq 1} \left(\sum_{dd'=n} \chi(d) d'^2 \right) q^n \quad G_2(q) = c_2 + \sum_{n \geq 1} \left(\sum_{d|n} \chi(d) d^2 \right) q^n$$

where the summations are taken over positive divisors d and d' . By the construction, $G_1(q)$ and $G_2(q)$ are independent elements of $M_3(\Gamma_1(4))$, thus they form a basis for the space over $\mathbb{C}$. Since θ_6 is also in this space, we can write it in terms of G_1 and G_2 , by solving for a and b in $\theta_6 = aG_1 + bG_2$. We do so by computing the first few coefficients of each modular form.

$$\begin{aligned}
\theta_6(q) &= 1 + 12q + 60q^2 + 160q^3 + 252q^4 + \mathcal{O}(q^5) \\
G_1(q) &= c_1 + q + 4q^2 + 8q^3 + 16q^4 + 26q^5 + \mathcal{O}(q^6) \\
G_2(q) &= c_2 + q + q^2 - 8q^3 + q^4 + 26q^5 + \mathcal{O}(q^6)
\end{aligned}$$

From here, we calculate that $a = 16$ and $b = -4$, hence we conclude

$$\begin{aligned}
\theta_6(q) &= 1 + \sum_{n \geq 1} \left[16 \left(\sum_{dd'=n} \chi(d) d'^2 \right) - 4 \left(\sum_{d|n} \chi(d) d^2 \right) \right] q^n \\
r_{\mathbb{Z}^6}(m) &= 16 \left(\sum_{dd'=m} \chi(d) d'^2 \right) - 4 \left(\sum_{d|m} \chi(d) d^2 \right)
\end{aligned}$$