MODULAR FORMS AND APPLICATIONS IN NUMBER THEORY

ALEX MINE

Abstract. Modular forms are complex analytic objects, but they also have many intimate connections with number theory. This paper introduces some of the basic results on modular forms, and explores some of their uses in number theory.

Contents

1. Modular Forms

1.1. Definitions.

Definition 1.1. The *upper half-plane*, denoted $\mathcal{H}$, is the set $\{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}$.

Definition 1.2. The *modular group* is another name for the group $\mathrm{SL}_2(\mathbb{Z})$, which is the group of 2×2 matrices with integer entries and determinant 1.

Throughout this paper, the modular group will act on $\mathcal{H}$ via the following: for $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$, $\tau \in \mathcal{H}$, define $\gamma(\tau) = \frac{a\tau+b}{c\tau+d}$. Note that if $\tau = x + iy$, then

$$\gamma(\tau) = \frac{(ax+b)+iay}{(cx+d)+icy} = \frac{((ax+b)+iay)((cx+d)-icy)}{|c\tau+d|^2}$$

which, by expanding, gives

$$\mathrm{Im}(\gamma(\tau)) = \frac{ady-bcy}{|c\tau+d|^2} = \frac{\mathrm{Im}(\tau)}{|c\tau+d|^2}. \tag{1.3}$$

Therefore, the action of $\mathrm{SL}_2(\mathbb{Z})$ fixes $\mathcal{H}$. Note that negating all of the entries of a given element of $\mathrm{SL}_2(\mathbb{Z})$ does not affect the action of the element.

Definition 1.4. A *modular form* is a function f on $\mathcal{H}$ satisfying:

(1) for all $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$, $\tau \in \mathcal{H}$, we have $f(\gamma(\tau)) = (c\tau + d)^k f(\tau)$ for some integer k, called the *weight* of f.
(2) f is holomorphic on $\mathcal{H}$
(3) f is holomorphic at ∞

By "holomorphic at ∞", we mean the following: let $q = e^{2\pi i\tau}$ and define the function $g(q) = f(\frac{\log q}{2\pi i})$. This is a well-defined function despite the log, because by letting $\gamma = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ in condition (1), we see that $f(\tau + 1) = f(\tau)$, so f takes the same value on each of the multiple arguments given by the logarithm. We will think of ∞ as lying infinitely far in the imaginary direction, and $q \to 0$ as $\mathrm{Im}(\tau) \to \infty$, so we say that f is holomorphic at ∞ if and only if g extends holomorphically to 0.

The modular forms of a given weight k form a vector space, denoted $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$.

1.2. **Connection to number theory.** The way that we will use modular forms to investigate number theory is through their Fourier expansions or *q-expansions*, so called because of the notation $q = e^{2\pi i\tau}$. The existence of such expansions is given by the condition of holomorphy at ∞. One important example of a modular form is the *Eisenstein series*:

$$G_k(\tau) = \sum_{(c,d)} \frac{1}{(c\tau + d)^k}$$

where the sum extends over all $(c, d) \in \mathbb{Z}^2 \setminus (0, 0)$. One can show that for even $k > 2$, the series converges uniformly and

$$G_k(\tau) = 2\zeta(k) + 2\frac{(2\pi i)^k}{(k-1)!} \sum_{n=1}^{\infty} \sigma_{k-1}(n) q^n$$

where ζ is the Reimann zeta function and $\sigma_a(n) = \sum_{d|n} d^a$, where the sum extends over the positive divisors of n.

We also define the *normalized Eisenstein series*, denoted $E_k(\tau)$, by $G_k(\tau)/(2\zeta(k))$. This is normalized in the sense that its constant term is always 1. It turns out that $M_8(\mathrm{SL}_2(\mathbb{Z}))$ is a 1-dimensional vector space, and thus E_4^2 and E_8, which are two of its elements, must be equal, as they have equal constant terms. Using this fact, we can equate the rest of the coefficients in their respective series. After expanding out the series for E_4^2, we arrive at the unexpected fact that for all integers $n \geq 1$,

$$\sigma_7(n) = \sigma_3(n) + 120 \sum_{k=1}^{n-1} \sigma_3(k)\sigma_3(n-k).$$

Similar methods will be used to investigate other problems below.

1.3. **The fundamental domain.** A central object relating to the modular group $\mathrm{SL}_2(\mathbb{Z})$ is its *fundamental domain*. This is a subset $\mathcal{D}$ of $\mathcal{H}$ with the property that for every element of $\mathcal{H}$, there is exactly one (in most cases) $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ such that $\gamma(\tau) \in \mathcal{D}$. The cannonical choice of $\mathcal{D}$ is $\{\tau \in \mathcal{H} : -\frac{1}{2} \leq \mathrm{Re}(\tau) \leq \frac{1}{2},\ |\tau| \geq 1\}$.

Proposition 1.5. *For $\mathcal{D}$ as defined above, every $\tau \in \mathcal{H}$ is transformed into $\mathcal{D}$ by some $\gamma \in SL_2(\mathbb{Z})$*

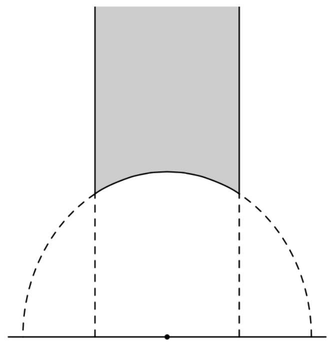

FIGURE 1. The fundamental domain $\mathcal{D}$.

Proof. The element $\begin{pmatrix} 1 & \pm 1 \\ 0 & 1 \end{pmatrix}$ of $\mathrm{SL}_2(\mathbb{Z})$ takes τ to $\tau \pm 1$, so repeatedly applying it will take any element to the strip $\{|\mathrm{Re}(z)| \leq \frac{1}{2}\}$, because the strip has length 1. Let τ' be the transformation of τ in this way. Then $\tau' \in \mathcal{D}$ unless $|\tau'| < 1$. If this is the case, then $\mathrm{Im}(-\frac{1}{\tau'}) = \mathrm{Im}(\frac{\tau'}{|\tau'|^2}) > \mathrm{Im}(\tau')$. Therefore we can apply the transformation $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tau' = -\frac{1}{\tau'}$, and then use the previous transformation to move the resulting point back into the strip. Each time we do this, the imaginary part increases, as long as the result still has magnitude less than 1. However, because there are only finitely many points on the lattice generated by 1 and τ' that lie within a given disk, there are only finitely many integers c, d such that $|c\tau' + d| < 1$. Combined with the fact established above that $\mathrm{Im}(\gamma(\tau)) = \frac{\mathrm{Im}(\tau)}{|c\tau+d|^2}$, this shows that the imaginary part of τ' can only take on finitely many values that are above its original value. Therefore, the second step will eventually no longer result in a point with greater imaginary part, which means that the resultant point will have magnitude greater than 1. Therefore the repeated application of the two transformations eventually results in a point inside of $\mathcal{D}$. □

It is clear, however, that for some τ, the element $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ that sends τ into $\mathcal{D}$ is not unique. For instance, if $\mathrm{Re}(\tau) = k + \frac{1}{2}$, with $k \in \mathbb{Z}$, then both $\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & k-1 \\ 0 & 1 \end{pmatrix}$ take τ into $\mathcal{D}$. However, this case and another are the only two possible ways that there can be any ambiguity. To state this fact differently, suppose $\tau \in \mathcal{H}$ and $\gamma_1, \gamma_2 \in \mathrm{SL}_2(\mathbb{Z})$, and both $\gamma_1(\tau), \gamma_2(\tau) \in \mathcal{D}$. Let $\tau_1 = \gamma_1(\tau)$ and $\tau_2 = \gamma_2(\tau)$. Then $\gamma_2\gamma_1^{-1}(\tau_1) = \tau_2$. Thus we have $\tau_1, \tau_2 \in \mathcal{D}$ and $\gamma = \gamma_2\gamma_1^{-1}$ such that $\gamma(\tau_1) = \tau_2$. The next result enumerates the cases in which this is possible.

Proposition 1.6. *Suppose $\tau \in \mathcal{D}$ and $\gamma \in SL_2(\mathbb{Z})$ such that $\gamma(\tau) \in \mathcal{D}$, with $\gamma \neq \pm I$. Then one of the following is true:*

(1) $\gamma = \pm \begin{pmatrix} 1 & \pm 1 \\ 0 & 1 \end{pmatrix}$

(2) $\gamma = \pm \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$

Proof. Without loss of generality, assume $\mathrm{Im}(\gamma(\tau)) \geq \mathrm{Im}(\tau)$ (otherwise invert γ and swap τ with $\gamma(\tau)$). Let $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. By equation (1.3), we thus have $|c\tau + d|^2 \leq 1$. Also, $\mathrm{Im}(\tau) \geq \frac{\sqrt{3}}{2}$ because this is true of any point in $\mathcal{D}$. Therefore

$$|c|\frac{\sqrt{3}}{2} \leq |c|\mathrm{Im}(\tau) = |\mathrm{Im}(c\tau + d)| \leq |c\tau + d| \leq 1.$$

Thus $c \in \{-1, 0, 1\}$. If $c = 0$, then $ad = 1$, so $a = d = \pm 1$, so $\gamma(\tau) = \tau_1 \pm b$. This can only be the case if τ is on the left or right boundary and $b = \pm 1$, which leads to (1).

Now suppose $c = \pm 1$. Then $|\tau \pm d|^2 \leq 1$, so

$$\begin{aligned} (\mathrm{Re}(\tau) \pm d)^2 + (\mathrm{Im}(\tau))^2 \leq 1 \quad &\Longrightarrow \quad (\mathrm{Re}(\tau) \pm d)^2 \leq 1 - (\mathrm{Im}(\tau))^2 \leq 1 - \tfrac{3}{4} = \tfrac{1}{4} \\ &\Longrightarrow \quad |\mathrm{Re}(\tau \pm d)| \leq \tfrac{1}{2} \end{aligned}$$

which means that $|d| \leq 1$, because $|\mathrm{Re}(\tau)| \leq \frac{1}{2}$. Now we have to examine the cases for d. If $d = 0$, then we in fact have $|\tau| \leq 1$, so $|\tau| = 1$ because $\tau \in \mathcal{D}$, and $\mathrm{Im}(\tau) = \mathrm{Im}(\gamma(\tau))$ by (1.3). But then we can apply the same analysis to $\tau' = \gamma(\tau)$ and $\gamma' = \gamma^{-1}$, because now τ' and γ' also satisfy the condition that $\mathrm{Im}(\gamma'(\tau')) \geq \mathrm{Im}(\tau')$. Thus we can conclude that $|\tau'| = 1$ as well. Then we have τ and $\gamma(\tau)$ both have magnitude 1 and have the same imaginary part but are distinct, so $\gamma(\tau) = -\frac{1}{\tau}$, from which (2) follows.

If $|d| = 1$, then we have $|\mathrm{Re}(\tau) \pm 1| \leq \frac{1}{2}$. However, because $|\mathrm{Re}(\tau)| \leq \frac{1}{2}$, we must have equality and $|\mathrm{Re}(\tau)| = \frac{1}{2}$. By the inequalities at the beginning of the previous paragraph, this gives $|\mathrm{Im}(\tau)| = \frac{\sqrt{3}}{2}$, so τ is one of the two corners of $\mathcal{D}$. In this case, both (1) and (2) work. □

The previous results establish an important fact: $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ generate $\mathrm{SL}_2(\mathbb{Z})$. Indeed, S and T are the same transformations used in Proposition 1.5, and they are (up to sign) the elements in the conclusion of Proposition 1.6. Therefore, we can pick $\tau \in \mathcal{D}$, and $\gamma \in \mathrm{SL}_2(\mathbb{Z})$. If $\gamma(\tau) \notin D$, then we can use proposition 1.5 to move $\gamma(\tau)$ back into $\mathcal{D}$ by using S and T. If the resulting point is τ, then we have written γ in terms of S and T. If not, we apply Proposition 1.6 and simply add on one more instance of $\pm T^{\pm 1}$ or $\pm S$. This last step also takes care of the case that $\gamma(\tau) \notin \mathcal{D}$. Note that $S^2 = -I$, so the sign differences can be dealt with.

1.4. **Congruence Subgroups.** In addition to $\mathrm{SL}_2(\mathbb{Z})$, it will also be useful to consider special subgroups of the same. We define the *principle congruence subgroup of level* N as

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

We will also write the congruences as $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Such subgroups are normal, because each is the kernel of the natural homomorphism taking $\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$. It is also true that each of these subgroups has finite index in $\mathrm{SL}_2(\mathbb{Z})$.

Definition 1.7. A *congruence subgroup of level N* is a subgroup $\Gamma \leq \mathrm{SL}_2(\mathbb{Z})$ such that $\Gamma(N) \leq \Gamma$.

Because $\Gamma(N)$ has finite index, any congruence subgroup does as well.

Two important classes of congruences subgroups are the following:

$$\Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \pmod N \right\}$$

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

where the congruence is entrywise, and a $*$ indicates that there are no restrictions on that entry (besides those imposed by membership of $\mathrm{SL}_2(\mathbb{Z})$).

We can define weight-k modular forms with respect to a congruence subgroup Γ in a similar way as modular forms above. In condition (1), we only check $\gamma \in \Gamma$, and we also require that $f(\alpha(\tau))(c\tau + d)^{-k}$ be holomorphic at ∞ for each $\alpha \in \mathrm{SL}_2(\mathbb{Z})$, where c and d make up the bottom row of α. We denote the vector space of modular forms of weight k with respect to Γ as $\mathcal{M}_k(\Gamma)$.

Congruence subgroups also have fundamental domains. In fact, their fundamental domains consist of unions of translates of $\mathcal{D}$ by elements of $\mathrm{SL}_2(\mathbb{Z})$, where each translating element is taken from one coset of the congruence subgroup in $\mathrm{SL}_2(\mathbb{Z})$. Thus the number of regions comprising the fundamental domain of Γ is the index of Γ in $\mathrm{SL}_2(\mathbb{Z})$. This will be unproven, but some intuition may be provided by the next proposition.

Proposition 1.8. *Let Γ be a congruence subgroup and suppose $\bigcup_j \{\pm I\}\Gamma\gamma_j = SL_2(\mathbb{Z})$. Then for every $\tau \in \mathcal{H}$, the orbit $\Gamma\tau$ is equal to an orbit $\Gamma\tau'$ for some $\tau' \in \bigcup_j \gamma_j \mathcal{D}$.*

Proof. We know that for each $\gamma \in \mathrm{SL}_2(\mathbb{Z})$, there is some $\jmath$ and some $\gamma' \in \Gamma$ such that $\pm I\gamma'\gamma_j = \gamma$. Thus $\mathrm{SL}_2(\mathbb{Z})$ is the union of all such elements, because any element of that form is clearly in $\mathrm{SL}_2(\mathbb{Z})$. However, the union of all of the elements of a group inverted is equal to the original group, so $\mathrm{SL}_2(\mathbb{Z})$ is also the union of all elements of the form $(\pm I\gamma'\gamma_j)^{-1} = \gamma_j^{-1} \pm I\gamma'^{-1}$. However, γ'^{-1} runs through $\mathrm{SL}_2(\mathbb{Z})$ as γ' does, so the union of such elements is $\bigcup_j \gamma_j^{-1}\{\pm I\}\Gamma$, and thus this union is equal to $\mathrm{SL}_2(\mathbb{Z})$. Next, pick $\tau \in \mathcal{H}$. Then there is some $\gamma' \in \Gamma$ and some j such that $\gamma_j^{-1} \pm I\gamma'(\tau) \in \mathcal{D}$. Then $\gamma(\tau) \in \gamma_j\mathcal{D}$ (we can drop the $\pm I$ because it does not affect the transformation). Choosing $\tau' = \gamma(\tau)$, we then see that $\Gamma\tau = \Gamma\tau'$ and $\tau' \in \gamma_j\mathcal{D} \in \bigcup_j \gamma_j\mathcal{D}$. This completes the proof. $\square$

2. A theorem

We will now prove an important theorem in the study of modular forms.

Theorem 2.1. *Let $f \in \mathcal{M}_k(SL_2(\mathbb{Z}))$. Then*

$$v_\infty(f) + \frac{1}{3}v_\rho(f) + \frac{1}{2}v_i(f) + \sum_{\substack{P \in \mathcal{D} \\ P \neq i,\rho}} v_P(f) = \frac{k}{12}$$

where $v_p(f)$ is the order of f at point p and $\rho = e^{\frac{2}{3}\pi i}$.

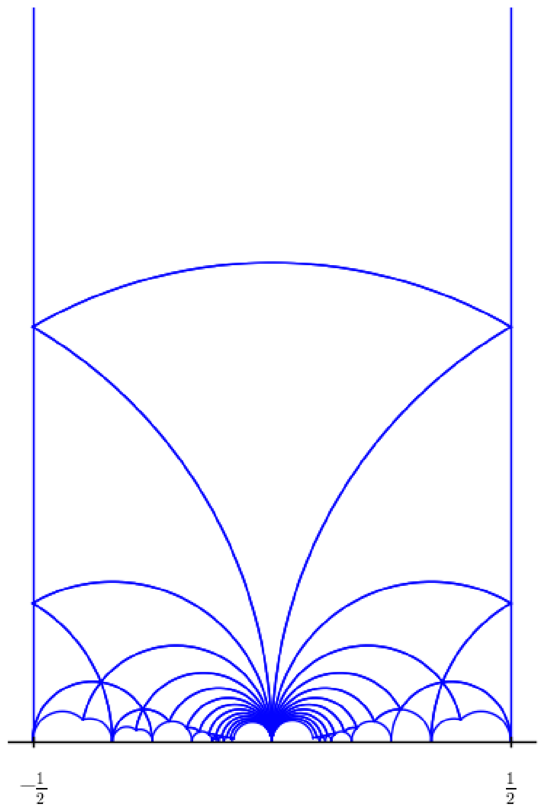

FIGURE 2. The fundamental domain of $\Gamma_0(24)$, shown as a union of translates of $\mathcal{D}$.

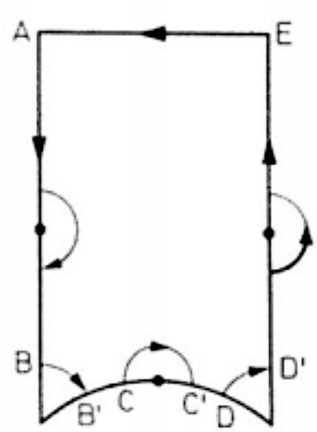

FIGURE 3. The path of integration in Theorem 2.1.

Proof. We will prove the theorem by integrating f around the contour shown in Figure 3.

If f has any poles or zeroes along the sides this contour, we can place half circles around them, as shown in Figure 3 between A and B and between D' and E. Placing these bumps in this way will keep the point inside the region and by the periodicity of f, the integrals along the bumps will cancel out. Similarly, if any poles or zeroes exist on the curved portion, we can place circular arcs around the point and the integrals will cancel out as the radii tend to 0. Therefore we can assume that f has no zeroes or poles on the boundary besides at i or ρ. The contour that we chose avoids these points. Then by the residue theorem, we have $\frac{1}{2\pi i}\oint \frac{f'}{f}dz = \sum \mathrm{Res}[f,p] = \sum_{\substack{P\in\mathcal{D}\\ P\neq i,\rho}} v_P(f)$ as the region approximates $\mathcal{D}$. Now we'll compute the integral over all of the components of the path. By changing variables to q, the top segment becomes a circle around infinity. Therefore, as the top segment goes to ∞, the value of the integral over it goes to $v_\infty(f)$. Next, by the periodicity of f, the integrals along the left and right sides cancel each other out. Next we look at the arc BB'. Near ρ, we can write $f = a(z-\rho)^m(1+\cdots)$, so $f'/f = m/(z-\rho) +$ holomorphic terms. The holomorphic terms have 0 integral as the radius of the arc shrinks, so we are left with integrating m/z over a arc approaching an angle of $\pi/3$, so the integral tends to $-m/6$. The arc around $-\bar{\rho}$ gives the same value, so in total we get $-\frac{1}{3}v_\rho(f)$. The same argument shows that the small arc in the middle gives a contribution of $-\frac{1}{2}v_i(f)$.

Finally, we have the arcs $B'C$ and $C'D$. The element S maps $B'C$ to DC', and we know $f(Sz) = z^k f(z)$. We compute

$$\frac{df(Sz)}{dz} = f'(Sz)\frac{1}{z^2} = z^k f'(z) + 2kz^{k-1}f(z)$$

which we can rearrange to

$$\frac{1}{z^2}\frac{f'(Sz)}{f(Sz)} = \frac{f'(z)}{f(z)} + \frac{k}{z}$$

By change of variables, we have

$$\int_{C'}^{D} \frac{f'(z)}{f(z)}dz = \int_{C}^{B'} \frac{f'(Sz)}{f(Sz)}dz.$$

Therefore the two integrals cancel, except for a term $\frac{k}{12}\int_{B'}^{C}\frac{k}{z}dz$. This term approaches $k/12$ in a similar manner to the above integrals. Combining all of these integrals proves the theorem. □

This is a useful result, because it allows us to classify the elements of $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$. For example: if $k=2$, there is no way for the left hand side to equal $1/6$, so $\mathcal{M}_2(\mathrm{SL}_2(\mathbb{Z}))$ is empty. If $k=4$, then the right hand side is $1/3$, so we must have $v_\rho(f) = 1$ and there must not be any other zeroes, unless f is identically 0. We know G_4 is in $\mathcal{M}_4(\mathrm{SL}_2(\mathbb{Z}))$, so $G_4(\rho) = 0$. Then there is some a such that $f - aG_4$ has a zero at ∞ and thus is identically 0, so $f = aG_4$. Thus $\mathcal{M}_4(\mathrm{SL}_2(\mathbb{Z}))$ is a one-dimensional vector space generated by G_4, denoted $\mathcal{M}_4 = (G_4)$. Similarly, we can show that $\mathcal{M}_6 = (G_6)$, $\mathcal{M}_8 = (G_4^2)$, $\mathcal{M}_10 = (G_4G_6)$, and in fact that every element of $\mathcal{M}(\mathrm{SL}_2(\mathbb{Z})) = \bigcup_k \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ is a polynomial in G_4 and G_6. Note that we have demonstrated the claim in section 1.2.

3. The k-Squares Problem

We now move on to investigate an important problem in number theory: in how many ways can one write an integer n as a sum of k squares? Let $r(n,k)$ be the number of such representations, where order does not matter. One property of this function that is easy to see is that if $a+b=k$, then $r(n,k)=\sum_{i=0}^{n} r(i,a)r(n-i,b)$. This is a Cauchy product, and suggests that we construct power series with these functions as coefficents. We thus define the generating functions

$$\theta(\tau,k)=\sum_{n=0}^{\infty} r(n,k)q^n.$$

It turns out that such series are absolutely convergent, and by our observation above, we have $\theta(\tau,k_1)\theta(\tau,k_2)=\theta(\tau,k_1+k_2)$. Because $q=e^{2\pi i\tau}$, we have $\theta(\tau+1,k)=\theta(\tau,k)$. One can use these θ-functions to investigate the values of $r(n,k)$; in particular, we will derive the values of $r(n,4)$.

Theorem 3.1. *For all integers $n>0$, $r(n,4)=8\sum_{\substack{d\mid n\\ 4\nmid d}} d$.*

Proof. Now let $\theta(\tau)=\theta(\tau,1)$. Then by definition, $\theta(\tau)=\sum_{d\in\mathbb{Z}} q^{d^2\tau}$, because the only numbers that can be written as the sum of one square are the squares themselves, and it can be done with two numbers, one positive and one negative (except for 0). One can show that $\theta(\frac{\tau}{4\tau+1})=\sqrt{4\tau+1}\theta(\tau)$ and thus $\theta(\frac{\tau}{4\tau+1},4)=(4\tau+1)^2\theta(\tau,4)$. Thus θ^4 is a weight-2 modular form with respect to the subgroup generated by $\pm\begin{pmatrix}1&1\\0&1\end{pmatrix}$ and $\pm\begin{pmatrix}1&0\\4&1\end{pmatrix}$, because these are the elements which θ^4 respects with regard to the modularity condition. This subgroup turns out to be $\Gamma_0(4)$. Note that $\theta^4(\tau)=\theta(\tau,4)$. Therefore by studying θ^4 we will be able to find the values of $r(n,4)$.

We know that $G_2(\tau)=2\zeta(2)-8\pi^2\sum_{n=1}^{\infty}\sigma(n)q^n$, but the series defining G_2 is only conditionally convergent, so it is not a modular form. However, it turns out that if we define

$$G_{2,N}(\tau)=G_2(\tau)-NG_2(N\tau),$$

then $G_{2,N}\in\mathcal{M}_2(\Gamma_0(N))$. We will examine the functions $G_{2,2}$ and $G_{2,4}$. First we will find explicit expressions for their q-coefficients.

We have

$$\begin{aligned}
G_{2,2}(\tau)&=G_2(\tau)-2G_2(2\tau)\\
&=2\zeta(2)-8\pi^2\sum_{n=1}^{\infty}\sigma(n)q^n-2(2\zeta(2)-8\pi^2\sum_{n=1}^{\infty}\sigma(n)q^{2n})\\
&=-2\zeta(2)+8\pi^2\sum_{n-1}^{\infty}\sigma^{\nmid 2}(n)q^n\\
&=-\frac{\pi^2}{3}\left(1+24\sum_{n-1}^{\infty}\sigma^{\nmid 2}(n)q^n\right)
\end{aligned}$$

where $\sigma^{\nmid 2}(n)$ is the sum of the divisors of n that are not divisible by 2. Similarly, we find that

$$G_{2,4}(\tau)=-\pi^2\left(1+8\sum_{n=1}^{\infty}\sigma^{\nmid 4}(n)q^n\right).$$

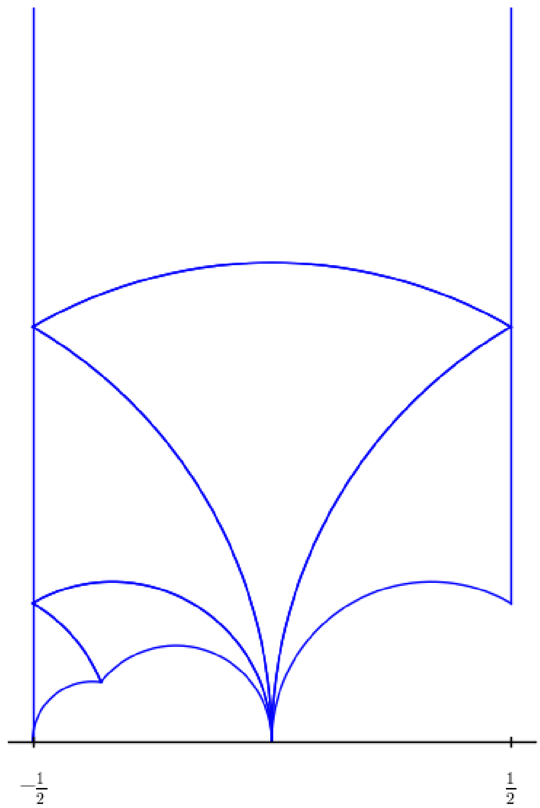

FIGURE 4. The fundamental domain for $\Gamma_0(4)$.

We will now use the normalized versions of these series, $E_{2,2} = G_{2,2} \cdot -\frac{3}{\pi^2}$ and $E_{2,4} = G_{2,4}/(-\pi^2)$.

Computing the first few terms of each, we get $E_{2,2} = 1 + 24q + \cdots$ and $E_{2,4} = 1 + 8q + \cdots$. Thus the two are linearly independent.

We can find a similar formula to that of Theorem 2.1 by integrating around the fundamental domain of $\Gamma_0(4)$. In this case, the formula turns out to be

$$\sum_{P \in X_0(4)} v_P(f) = \frac{k}{2}$$

where $X_0(4)$ denotes the fundamental domain of $\Gamma_0(4)$ under some suitable identifications along the boundary, similar to how the two vertical edges of $\mathcal{D}$ can be identified because they are equal under $\mathrm{SL}_2(\mathbb{Z})$. Now, for each P, define $f_P = aE_{2,2} + bE_{2,4}$ with a and b chosen such that $f_P(P) = 0$. Then if $g \in \mathcal{M}_2(\Gamma_0(4))$ and $g(P) = 0$, $g/f_P \in \mathcal{M}_0$ so it is a constant complex number. Thus every element of $\mathcal{M}_2$ is a linear combination of $E_{2,2}$ and $E_{2,4}$.

Now expand out θ^4 to get $1 + 8q + \cdots$. From this we can see that $\theta^4 = E_{2,4}$ because the first two coefficients match and we have established that the vector space is two-dimensional so that's all we need. Thus we can equate the coefficients and find that $r(n, 4) = 8\sigma^{\not 4}(n)$ as desired. □

Thus we have used the theory of modular forms to solve a problem in number theory. Other values of k can be attacked in similar ways.

Acknowledgments. I would like to thank my mentor, Drew Moore, for introducing me to this material and presenting it in a coherent and interesting ways, along with his help and guidance through the process of creating this paper.

References

[1] Fred Diamond and Jerry Shurman. A First Course in Modular Forms. Springer Science and Business Media, Inc. 2005.
[2] Serge Lang. Introduction to Modular Forms. Springer-Verlag Berlin Heidelberg. 1976.
[3] The LMFBD Collaboration. The L-functions and Modular Forms Database. http://www.lmfdb.org