MODULAR FORMS

MARIGOLD STRUPP

1. Introduction

Modular forms are holomorphic functions on the upper half-plane that satisfy a strong symmetry under the action of $SL_2(\mathbb{Z})$. This symmetry severely restricts their behavior: a modular form is determined by its values on a small region, and its Fourier expansion exhibits a highly structured form with significant arithmetic content.

One way to understand the origin of this symmetry is through lattices in $\mathbb{C}$. Writing a lattice in the form $\mathbb{Z} + z\mathbb{Z}$, a change of basis corresponds to the action of $SL_2(\mathbb{Z})$ on z. It is therefore natural to study functions whose transformation under this action is tightly controlled. This perspective leads directly to the definition of modular forms and explains the form of the modularity condition.

In this paper, we introduce modular forms from first principles. We begin by defining the action of $SL_2(\mathbb{Z})$ and the notion of modularity. We then study Fourier expansions and construct Eisenstein series as basic examples. Next, we describe the structure of the graded algebra of modular forms, culminating in a proof that it is generated by two elements. Finally, we introduce Hecke operators and examine their effect on Fourier coefficients, with particular attention to eigenforms.

2. The definition of modular forms

Modular forms are a certain sort of extremely symmetric object, that is they have some amount of invariance under a certain group action. So, to begin with, we should define what group is acting and exactly how.

Definition 2.1. Define

$$\mathrm{SL}_2(\mathbb{Z})$$

as the set of all invertible 2×2 matrices M with integer entries and determinant 1. The requirement that the determinant be 1 ensures that if $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is contained in $\mathrm{SL}_2(\mathbb{Z})$ then its inverse $M^{-1} = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$ Is also contained in $\mathrm{SL}_2(\mathbb{Z})$. Two particularly important elements of this group are

$$S = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \qquad T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

Theorem 2.2. $\mathrm{SL}_2(\mathbb{Z})$ *is generated by* S *and* T

Proof. First note that

$$S^2 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

and hence S has order four. We also have

$$T^n = \begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix}$$

Now take any matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and split into two cases. First if $c = 0$ then because the determinant is 1 we must have $ad = 1$ so either $a = d = 1$ or $a = d = -1$. In the first case the matrix has the form

$$\begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix} = T^n$$

and in the second it has the form

$$\begin{bmatrix} -1 & n \\ 0 & -1 \end{bmatrix} = -T^{-n} = S^2T^{-n}$$

Now suppose $c \neq 0$. Then first we articulate the actions of S and T^n in general

$$S\begin{bmatrix} x & y \\ z & w \end{bmatrix} = \begin{bmatrix} -z & -w \\ x & y \end{bmatrix} \qquad T^n\begin{bmatrix} x & y \\ z & w \end{bmatrix} = \begin{bmatrix} x+zn & y+wn \\ z & w \end{bmatrix}$$

Now we want to, via these actions, reduce c to 0. To do so we produce a sequence of multiplications on the left by S and T which always strictly reduces the absolute value of the lower left corner. First write $a = qc + r$ where $0 \leq |r| < |c|$, then

$$S(T^{-q}\begin{bmatrix} a & b \\ c & d \end{bmatrix}) = S\begin{bmatrix} a-cq & b-qd \\ c & d \end{bmatrix} = \begin{bmatrix} -c & -d \\ r & b-qd \end{bmatrix}$$

which has lower left corner with $|r| < |c|$. Repeating this process finitely many times will then reduce this to a matrix with lower left corner equal to 0. Hence by multiplying on the left by finitely many S, T, and T^{-1}, we can produce an element of the form T^n or S^2T^n. Then we must have that $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is in the subgroup of $\mathrm{SL}_2(\mathbb{Z})$ generated by S and T, but this matrix is arbitrary so S and T generate $\mathrm{SL}_2(\mathbb{Z})$. ■

This group acts on what's known as the upper half plane

Definition 2.3. Define

$$\mathbb{H} = \{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}$$

Proposition 2.4. $\mathrm{SL}_2(\mathbb{Z})$ *acts on* $\mathbb{H}$ *by*

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} : z \mapsto \frac{az+b}{cz+d}$$

And in particular this action extends to a right action on functions $f : \mathbb{H} \to \mathbb{C}$ *given by*

$$f \mapsto f \circ \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Proof. We have

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} z = \frac{1(z)+0}{0(z)+1} = z$$

and

$$\begin{aligned}
\begin{bmatrix} a & b \\ c & d \end{bmatrix} \left(\begin{bmatrix} e & f \\ g & h \end{bmatrix} z \right) &= \begin{bmatrix} a & b \\ c & d \end{bmatrix} \frac{ez+f}{gz+h} \\
&= \frac{a(\frac{ez+f}{gz+h}) + b}{c(\frac{ez+f}{gz+h}) + d} \\
&= \frac{aez + af + bgz + bh}{cez + cf + dgz + dh} \\
&= \frac{(ae+bg)z + af + bh}{(ce+dg)z + cf + dh} \\
&= \begin{bmatrix} ae+bg & af+bh \\ ce+dg & cf+dh \end{bmatrix} z \\
&= \left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} e & f \\ g & h \end{bmatrix} \right) z
\end{aligned}$$

So to check it's a group action we need only check that this sends elements of the upper half plane to elements of the upper half plane. We check that S and T do so and hence that the whole group does. We have

$$Sz = \frac{-1}{z} = \frac{-\bar{z}}{|z|^2}$$

which has imaginary part $\frac{1}{|z|^2}\mathrm{Im}(z) > 0$. We also have

$$Tz = 1 + z$$

which has imaginary part $\mathrm{Im}(z) > 0$ ■

This finally allows us to define modular forms.

Definition 2.5. A modular form of weight k is a function $f : \mathbb{H} \to \mathbb{C}$ fulfilling the following conditions

- f is holomorphic on $\mathbb{H}$
- For any element $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ of $\mathrm{SL}_2(\mathbb{Z})$ $f(\gamma z) = (cz+d)^k f(z)$ (this is known as the modularity condition)
- $f(z)$ remains bounded as $\mathrm{Im} z \to \infty$ (this is known as being holomorphic at infinity)

We write M_k for the space of all modular forms of weight k.

Proposition 2.6. *The space M_k forms a $\mathbb{C}$ vector space under pointwise addition and $\bigoplus M_k$ is a graded algebra.*

Proof. Let f and g be modular forms of weight k and let a be a complex number, then for any $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ we have

$$\begin{aligned}
(af+g)(\gamma z) &= af(\gamma z) + g(\gamma z) \\
&= a(cz+d)^k f(z) + (cz+d)^k g(z) \\
&= (cz+d)^k (af+g)(z)
\end{aligned}$$

Hence M_k is a $\mathbb{C}$ vector space. Now suppose f, g are modular forms of weight k and j respectively. Then

$$f(\gamma z)g(\gamma z) = (cz+d)^k(cz+d)^j f(z)g(z) = (cz+d)^{k+j}f(z)g(z)$$

so $f(z)g(z)$ is a modular form of weight $k+j$. Hence $\bigoplus M_k$ is a graded $\mathbb{C}$ algebra. ∎

Now it is important to simplify the modularity condition a bit. To do so we use the following lemma.

Lemma 2.7. *If f satisfies the modularity condition for two matrices $\gamma_1, \gamma_2 \in \mathrm{SL}_2(\mathbb{Z})$ it satisfies the modularity condition for their product. In particular, this tells us that if f satisfies the modularity condition for S and T it does so for all of* $\mathrm{SL}_2(\mathbb{Z})$.

Proof. Let $\gamma_1 = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $\gamma_2 = \begin{bmatrix} e & f \\ g & h \end{bmatrix}$ and suppose f satisfies the modularity condition for each of them. Then

$$f(\gamma_1\gamma_2 z) = (c\gamma_2 z + d)^k f(\gamma_2 z) = (c\gamma_2 z + d)^k (gz+h)^k f(z) = (ecz + gdz + fc + hd)^k f(z)$$

since we have

$$\gamma_1\gamma_2 \begin{bmatrix} ea+gb & fa+hb \\ ec+gd & fc+hd \end{bmatrix}$$

it follows that f satisfies the modularity condition for $\gamma_1\gamma_2$. ∎

Remark 2.8. The modularity condition now reduces down to checking S and T. For S we need to check

$$f(\frac{-1}{z}) = z^k f(z)$$

and for T we need to even more simply check

$$f(z+1) = f(z)$$

This means that every value of f is determined by its value on whats called the fundamental domain. The subset $\mathcal{F}$ of $\mathbb{H}$ given by

$$\{z \in \mathbb{H} : |z| \geq 1, |\mathrm{Re}(z)| \leq \frac{1}{2}\}$$

In particular a modular form f of any weight must be bounded on $\mathcal{F}$ due to the holomorphicity at infinity condition.

The final point to make here is that we generally only need to consider even weights $k \geq 0$.

Proposition 2.9. *If k is not an even integer, or if $k < 0$ the only modular form of weight k is* 0.

Proof. If k is not an even integer then suppose f is a modular form of weight k and consider the matrix $-I = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$. Then $-Iz = \frac{-z}{-1} = z$ so we must have

$$f(z) = f(-Iz) = (-1)^k f(z)$$

Since k is not an even integer $(-1)^k \neq 1$ and so $f = 0$.

Now if f satisfies the modularity condition for weight $k < 0$. Define

$$F(z) = |f(z)|(\mathrm{Im} z)^{k/2}$$

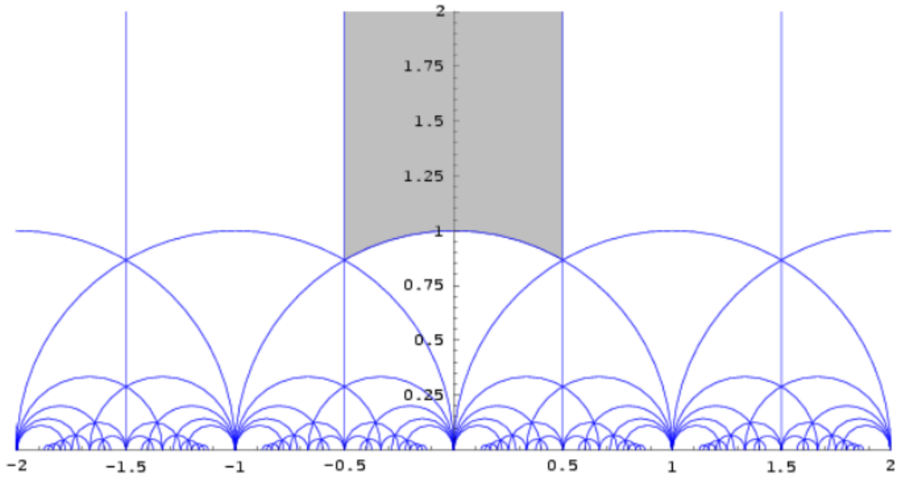

Figure 1. The fundamental domain

For $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ we have

$$f(\gamma z) = (cz+d)^k f(z), \qquad \mathrm{Im}(\gamma z) = \frac{\mathrm{Im} z}{|cz+d|^2}.$$

Taking absolute values and combining these identities gives

$$F(\gamma z) = F(z),$$

so F is invariant under the action of $SL_2(\mathbb{Z})$.

Because f is holomorphic at $i\infty$, there exists C such that $|f(x+iy)| \leq C$ for all y sufficiently large. Since $k < 0$, we have $y^{k/2} \to 0$ as $y \to \infty$, and therefore

$$F(x+iy) = |f(x+iy)|\, y^{k/2} \to 0 \qquad (y \to \infty).$$

Choose Y large enough that $|f(z)| \leq C$ whenever $\mathrm{Im} z \geq Y$, and put

$$\mathcal{F}_Y = \{z \in \mathcal{F} : \mathrm{Im} z \leq Y\}.$$

The set $\mathcal{F}_Y$ is compact, and F is continuous, so F attains a maximum $M \geq 0$ on $\mathcal{F}_Y$. Since F tends to 0 as $\mathrm{Im}(z) \to \infty$, this M is the global maximum of F on $\mathbb{H}$.

Suppose $M > 0$. Then there exists $z_0 \in \mathcal{F}_Y$ with $F(z_0) = M$. Write $y_0 = \mathrm{Im} z_0$. Since $M > 0$, we have $f(z_0) \neq 0$.

Because F attains a strict local maximum at z_0, there exists a neighbourhood U of z_0 and $\varepsilon > 0$ such that

$$F(z) \leq M - \varepsilon \qquad (z \in U \setminus \{z_0\}).$$

By continuity of the function $z \mapsto (\mathrm{Im} z)^{k/2}$, after possibly shrinking U we may assume that

$$|(\mathrm{Im} z)^{k/2} - y_0^{k/2}| < \delta \qquad (z \in U)$$

for some $\delta > 0$ with $y_0^{k/2} - \delta > 0$ and

$$\frac{M - \varepsilon}{y_0^{k/2} - \delta} < \frac{M}{y_0^{k/2}}.$$

Then for $z \in U \setminus \{z_0\}$ we obtain

$$|f(z)| = \frac{F(z)}{(\mathrm{Im} z)^{k/2}} \leq \frac{M-\varepsilon}{(\mathrm{Im} z)^{k/2}} \leq \frac{M-\varepsilon}{y_0^{k/2}-\delta} < \frac{M}{y_0^{k/2}} = |f(z_0)|.$$

Thus $|f|$ has a strict local maximum at z_0. The Maximum Modulus Principle therefore implies that f is constant on $\mathbb{H}$.

If $f \equiv c \neq 0$, the modular transformation law would give $c = (cz+d)^k c$ for all $\gamma \in SL_2(\mathbb{Z})$, which is impossible since $(cz+d)^k$ is not constant when $k \neq 0$. Hence $f \equiv 0$, contradicting $M > 0$. Therefore $M = 0$, so $F \equiv 0$ and consequently $f \equiv 0$. ∎

Remark 2.10. It is worth recording a different perspective on modular forms which will become important later. Recall that a complex elliptic curve is a torus $\mathbb{C}/\Lambda$ where $\Lambda = \omega_1\mathbb{Z} + \omega_2\mathbb{Z}$ is a lattice, two lattices Λ, Λ' give isomorphic elliptic curves if there is some complex number c such that $\Lambda = c\Lambda'$.

It is a reasonable question to ask what "functions of elliptic curves" are, and while this has an answer (namely elliptic functions) it isn't quite modular forms.

A modular form is instead a function on the space of elliptic curves that sort of transforms like a differential. Consider two lattices Λ and $\lambda\Lambda$ which give the isomorphic elliptic curves, the isomorphism $\mathbb{C}/\lambda\Lambda \to \mathbb{C}/\Lambda$ is then $z \mapsto \lambda^{-1}z$. Under this transformation the k-fold differential dz^k transforms to $\lambda^{-k}dz^k$. Hence, if we want modular forms to "look like" differentials on elliptic curves we would need them to be functions F from the space of lattices to $\mathbb{C}$ such that it fulfills a sort of weight k homogeneity condition

$$F(\lambda\Lambda) = \lambda^{-k}F(\Lambda)$$

To recover our original definition from this one given some F as above consider the function $f : \mathbb{H} \to \mathbb{C}$ given by

$$f(z) = F(\mathbb{Z} + z\mathbb{Z})$$

the condition on F then becomes exactly the modularity condition on f. If we change basis on the lattice $\mathbb{Z} + z\mathbb{Z} = \Lambda_z$ via $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ then $z \mapsto \gamma z$ and the lattice scales by $(cz+d)^{-1}$. Hence the condition on F tells us

$$f(\gamma z) = F((cz+d)^{-1}\Lambda_z) = (cz+d)^k F(\Lambda_z) = (cz+d)^k f(z)$$

Conversely, by an identical argument any weight k modular form f gives a function F on lattices with the weight k homogeneity condition via

$$F(\omega_1\mathbb{Z} + \omega_2\mathbb{Z}) = \omega_2^{-k} f(\frac{\omega_1}{\omega_2})$$

This perspective allows several constructions with modular forms that appear random to become quite natural. We will return to it when we define Hecke operators.

3. q expansions and Eisenstein series

Given a modular form f note that

$$f(1+z) = f(Tz) = f(z)$$

and hence f is one periodic. This allows us to take a fourier expansion. And this Fourier expansion has a particularly nice form.

Lemma 3.1. *For any modular form f, if $q = e^{2\pi iz}$ then the Fourier expansion of f*

$$f(z) = \sum_{n=-\infty}^{\infty} a_n q^n$$

has $a_n = 0$ for all $n < 0$. In other words we may write

$$f(z) = \sum_{n=0}^{\infty} a_n q^n$$

Proof. By definition of modular forms f remains bounded as $\text{Im}(z) \to \infty$. However if $n < 0$ then as $\text{Im}(z) \to \infty$

$$(e^{2\pi iz})^n \to \infty$$

so if any $a_n \neq 0$ for $n < 0$ then

$$|f(z)| = \left| \sum_{n=-\infty}^{\infty} a_n q^n \right| \to \infty$$

But this contradicts the boundedness of f as $\text{Im}(z) \to \infty$. ∎

Remark 3.2. Notably, as $\text{Im}(z) \to \infty$ then $q \to 0$, so for any modular form f the constant coefficient in its q-expansion is

$$\lim_{z \to i\infty} f(z)$$

Definition 3.3. A modular form is said to be a cusp form if the coefficient a_0 of its q-expansion is 0. Equivalently f is a cusp form if it goes to 0 as $\text{Im} z \to \infty$. We denote the space of cusp forms with weight k by S_k.

Our first example of a modular form will be as follows.

Definition 3.4. For even $k \geq 4$ the weight k Eisenstein series is defined as

$$G_k(z) = \sum_{(m,n) \in \mathbb{Z}^2 n\{(0,0)\}} \frac{1}{(mz+n)^k}$$

Lemma 3.5. *The series $G_k(z)$ converges absolutely and (locally) uniformly, Further it is a modular form of weight k.*

Proof. For any $N \geq 2$ we show that G_k converges uniformly and absolutely on

$$U_N = \{z \in \mathbb{H} : \text{Im} z > \frac{1}{N}, |\text{Re}(z)| \leq N\}$$

Then for any $z \in U_N$ and any $x \in \mathbb{R}$ if $|x| \leq 2N$ then because $\text{Im} z > \frac{1}{N}$ we have

$$|z + x| \geq |\text{Im}(z)| \geq \frac{1}{N}$$

And if $|x| \geq 2N$ we have

$$|z + x| \geq |x| - |\text{Re}(z)| \geq |x| - N \geq \frac{|x|}{2}$$

Therefore

$$|z + x| \geq \frac{1}{2N^2} \max(1, |x|)$$

and so

$$|mz+n|^{-k} = |m^{-k}||z+n/m|^{-k} \leq |m|^{-k}2N^{2k}\max(1,|n/m|)^{-k} = (2N^2)^k\max(|n|,|m|)^{-k}$$

Therefore

$$\begin{aligned}\sum_{(m,n)\in\mathbb{Z}^2n\{(0,0)\}}\frac{1}{|mz+n|^k} &\leq (2N^2)^k\sum_{(m,n)\in\mathbb{Z}^2n\{(0,0)\}}\frac{1}{\max(|m|,|n|)^k}\\ &= (2N^2)^k\sum_{a=1}^{\infty}a^{-k}|\{(m,n)\in\mathbb{Z}^2n\{(0,0)\} : \max(|m|,|n|)=a\}|\\ &= (2N^2)^k\sum_{a=1}^{\infty}8a^{1-k}\\ &= 8(2A^2)^k\zeta(k-1)\end{aligned}$$

Hence $G_k(z)$ converges absolutely and uniformly on U_N, and so it is holomorphic on $\mathbb{H}$.

Now as $\mathrm{Im}(z)\to\infty$ if $m\neq 0$ then $\frac{1}{(mz+n)^k}\to 0$ and if $m=0$ then $\frac{1}{(mz+n)^k}\to n^{-k}$ so because k is even

$$G_k(z)\to\sum_{n=1}^{\infty}2n^{-k} = 2\zeta(k)$$

Finally to check the modularity condition for G let $\gamma = \begin{bmatrix} a & b\\ c & d\end{bmatrix}\in\mathrm{SL}_2(\mathbb{Z})$ then

$$\begin{aligned}\sum_{(m,n)\in\mathbb{Z}^2n\{(0,0)\}}\frac{1}{(m(\gamma z)+n)^k} &= \sum_{(m,n)\in\mathbb{Z}^2n\{(0,0)\}}\frac{1}{(m\frac{az+b}{cz+d}+n)^k}\\ &= \sum_{(m,n)\in\mathbb{Z}^2n\{(0,0)\}}\frac{(cz+d)^k}{(ma+nc)z+mb+nd)^k}\\ &= (cz+d)^k\sum_{(m,n)\in\mathbb{Z}^2n\{(0,0)\}}\frac{1}{(mz+n)^k}\end{aligned}$$

■

We have just shown that one of the Fourier coefficients of G_k (namely the constant one) has remarkable arithmetic significance, being equal to $2\zeta(k)$. We will now show that the remaining coefficients have similar significance.

Theorem 3.6. *The q-expansion of G_k is*

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

where

$$\sigma_{k-1}(n) = \sum_{d|n}d^{k-1}$$

Proof. Note that

$$\pi i - 2\pi i\sum_{m=1}^{\infty}e^{2\pi izm} = \pi i\frac{e^{2\pi iz}+1}{e^{2\pi iz}-1} = \pi\cot(\pi z)$$

and recall the cotangent formula

$$\pi \cot \pi z = \frac{1}{z} + \sum_{d \in \mathbb{Z} n \{0\}}^{\infty} \frac{1}{z-d} - \frac{1}{d}$$

Then

$$\frac{1}{z} + \sum_{d \in \mathbb{Z} n \{0\}}^{\infty} \frac{1}{z+d} - \frac{1}{d} = \pi i - 2\pi i \sum_{m=1}^{\infty} e^{2\pi i z m} = \pi i - 2\pi i \sum_{m=1}^{\infty} q^m$$

Now the $k-1$th derivative of $\frac{1}{z+m}$ is $-\frac{(k-1)!}{(1+z)^k}$ and the $k-1$th derivative of the right hand side is easily

$$-(2\pi i)^k \sum_{m=1}^{\infty} m^{k-1} q^m$$

Thus we have

$$-\sum_{m \in \mathbb{Z}} \frac{(k-1)!}{(z+m)^k} = -(2\pi i)^k \sum_{m=1}^{\infty} m^{k-1} q^m$$

and hence

$$\sum_{n \in \mathbb{Z}} \frac{1}{(z+n)^k} = \frac{(2\pi i)^k}{(k-1)!} \sum_{m=1}^{\infty} m^{k-1} q^m$$

substitute $z = rz$ to obtain

$$\sum_{n \in \mathbb{Z}} \frac{1}{(rz+n)^k} = \frac{(2\pi i)^k}{(k-1)!} \sum_{m=1}^{\infty} m^{k-1} q^{rm}$$

Now we have

$$\begin{aligned} G_k(z) &= \sum_{(r,n) \in \mathbb{Z}^2 n \{(0,0)\}} \frac{1}{(rz+n)^k} \\ &= 2\zeta(k) + 2 \sum_{r=1}^{\infty} \sum_{n \in \mathbb{Z}} \frac{1}{(rz+n)^k} \\ &= 2\zeta(k) + 2 \frac{(2\pi i)^k}{(k-1)!} \sum_{r=1}^{\infty} \sum_{m=1}^{\infty} m^{k-1} q^{rm} \\ &= 2\zeta(k) + 2 \frac{(2\pi i)^k}{(k-1)!} \sum_{m=1}^{\infty} \frac{m^{k-1} q^m}{1-q^m} \\ &= 2\zeta(k) + 2 \frac{(2\pi i)^k}{(k-1)!} \sum_{n=1}^{\infty} \sigma_{k-1}(n) q^n \end{aligned}$$

■

Definition 3.7. It is typically easier to work with normalized modular forms, that is those with constant coefficient 1 and so we define the normalized Eisenstein series.

$$E_k(z) = \frac{G_k(z)}{2\zeta(k)} = 1 - \frac{2k}{B_k} \sum_{n=1}^{\infty} \sigma_{k-1}(n) q^n$$

(The q expansion comes from writing $\zeta(k) = \frac{(-1)^{k/2-1}B_k(2\pi)^k}{2(k)!}$ for even k, B_k is the kth Bernouli number)

Remark 3.8. The above theorem works not only for those G_k which are modular forms, but in fact for the one that isn't, being G_2. In particular as we go remember that

$$E_2(z) = 1 - 24\sum_{n=1}^{\infty} \sigma_1(n)q^n$$

4. The Structure of $\bigoplus M_k$

The goal of this section is to give an explicit easy to work with form for the algebra

$$\bigoplus M_k$$

First we show that there is a weight 12 cusp form Which vanishes nowhere on $\mathbb{H}$.

Definition 4.1. Define the dedekind eta function

$$\eta(z) = q^{1/24}\prod_{n=1}^{\infty}(1-q^n)$$

and the ramanujan discriminant function.

$$\Delta_r(z) = \eta(z)^{24}$$

Theorem 4.2. *The ramanujan discriminant is holomorphic, non-vanishing on* $\mathbb{H}$*, and a cusp form of weight* 12*.*

Proof. We prove this via examining the eta function. Note that for $\operatorname{Im}(z) > 0$ we have $|q| < 1$. First of all to show it is holomorphic we need only show that its logarithm is absolutely convergent on compact sets. To do so note that

$$\log\eta(z) = \frac{\pi i z}{12} + \sum_{n=1}^{\infty}\log(1-q^n)$$

Now use the taylor expansion for $\log(1-x)$ to obtain

$$\left|\sum_{n=1}^{\infty}\log(1-q^n)\right| \leq \sum_{n=1}^{\infty}\sum_{j=1}^{\infty}\frac{|q|^{nj}}{j} = \sum_{n=1}^{\infty}\sigma_{-1}(n)|q|^n < \sum_{n=1}^{\infty} n|q|^n$$

because $0 < |q| < 1$ absolute uniform convergence follows.

To show its nonvanishing note that in order to vanish some $1 - q^n$ would have to vanish, but this is impossible as $0 < |q| < 1$.

We now show that

$$\eta(z+1) = e^{\pi i/12}\eta(z) \qquad \eta(\frac{-1}{z}) = \eta(z)\sqrt{\frac{z}{i}}$$

The first identity is relatively easy. Recall that $q = e^{2\pi i z}$ and so when $z \to z+1$ $q \to e^{2\pi i(z+1)} = e^{2\pi i} = q$ and $q^{1/24} \to e^{\pi i/12}q^{1/24}$ so

$$\eta(z+1) = e^{\pi i/12}\eta(z)$$

This second identity is significantly more cumbersome to prove. The following proof is fairly short but involves a number of technical manipulations. If some of the details feel difficult to follow, that is completely normal; the argument itself is not central to what follows, and the main point is the resulting transformation law.

As both sides of the identity are analytic we need only prove their logarithmic derivatives agree. Hence we need only show

$$\frac{\eta'(-1/z)}{z^2\eta(-1/z)} = \frac{\eta'(z)}{\eta(z)} + \frac{1}{2z}$$

Now define

$$\zeta(s;z) = \frac{1}{s} + \sum_{(m,n)\in\mathbb{Z}^2 n\{(0,0)\}} \frac{1}{s+mz+n} - \frac{1}{mz+n} + \frac{s}{(mz+n)^2}$$

And look at the two specific functions

$$\phi_1(z) = \zeta(1/2; z)$$

and

$$\phi_2(z) = \zeta(z/2; s)$$

Then the definition of $\zeta(s;z)$ immediately gives

$$\phi_2(z) = \frac{1}{z}\phi_1(\frac{-1}{z})$$

Using the cotangent sum as well as a similar formula for $\csc^2$ $\zeta(s;z)$ becomes

$$\zeta(s;z) = \frac{\pi^2}{3}s + \pi\cot\pi s + \sum_{0\neq m\in\mathbb{Z}} \pi(\cot\pi(mz+s) - \pi\cot m\pi z + \pi^2 s\csc^2\pi mr$$

then

$$\phi_1(z) = \frac{\pi^2}{6} + \pi^2\sum_{m=1}^{\infty}\csc^2\pi mz$$

and

$$\phi_2(z) = \frac{1}{2} + \frac{z\pi}{6} + \sum_{m=1}^{\infty}\csc^2\pi mz = z\phi_1(z) + \frac{1}{2}$$

Hence

$$\frac{1}{z}\phi_1(\frac{-1}{z}) = z\phi_1(z) + \frac{1}{2}$$

Now the q-expansion of $\csc^2\pi z$ is $-4\sum_{n=1}^{\infty} nq^n$ and the q-expansion of the logarithmic derivative of η is

$$\frac{d}{dz}\left(\frac{\pi iz}{12} + \sum_{n=1}^{\infty}\log(1-q^n)\right) = 2\pi i(1/24 - \sum_{n=1}^{\infty}\frac{nq^n}{1-q^n}) = 2\pi i(1/24 - \sum_{n,k\geq 1} nq^{nk})$$

Hence the q-expansion of ϕ_1 is

$$-2\pi i(\frac{\eta'(z)}{\eta})$$

So the transformation law we found of ϕ_1 becomes the desired identity for η.

Now raising η to the 24 power we obtain the following transformation laws for the ramanujan discriminant

$$\Delta_r(z+1) = e^{2\pi i}\Delta_r(z) = \Delta_r(z) \qquad \Delta_r(\frac{-1}{z}) = \Delta_r(z)\frac{z^{12}}{i^{12}} = z^{12}\Delta_r(z)$$

Therefore Δ_r fulfills the modularity condition.

Now because $q \to 0$ as $z \to i\infty$ it follows that

$$\Delta_r(z) = q \prod_{n=1}^{\infty}(1-q^n)^{24} \to 0$$

and hence is a cusp form.

■

That one was long but we get something truly awesome because of it.

Lemma 4.3. *For any modular form f of weight k with constant term a_0 the function*

$$\frac{f(z) - a_0 E_k(z)}{\Delta_r(z)}$$

is a modular form of weight $k-12$.

Proof. Because Δ_r is non-vanishing this function is holomorphic. As for the modularity condition $f(z) - a_0 E_k(z)$ is a modular form of weight k and so for any $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ we find

$$\frac{f(\gamma z) - a_0 E_k(\gamma z)}{\Delta_r(\gamma z)} = \frac{(cz+d)^k}{(cz+d)^{12}} \frac{f(z) - a_0 E_k(z)}{\Delta_r(z)} = (cz+d)^{k-12} \frac{f(z) - a_0 E_k(z)}{\Delta_r(z)}$$

■

Corollary 4.4. *For weights $0, 2, 4, 6, 8, 10$ M_k has dimension 1 except for $k=2$ when it has dimension 0*

Proof. First we show they all have dimension at most one. Because the only modular form of weight $k-12$ for these k is 0 it follows that for any weight k form f

$$\frac{f(z) - a_0 E_k}{\Delta_r} = 0$$

and hence because Δ_r is non-vanishing that

$$f = a_0 E_k$$

. Hence the dimension is at most one.

For $k \geq 4$ the Eisentstein series lies in M_k and for $k = 0$ constants lie in M_0 so each of these have dimension 1.

As for M_2 suppose f is in M_2. Then $f(\frac{-1}{z}) = z^2 f(z)$ and thus $f(i) = -i$ so $f(i) = 0$. Now $f(z)^2$ is a modular form of weight 4 and as we have just shown, is therefore a multiple of E_4. write $f^2 = cE_4$, we have

$$0 = f(i)^2 = c\left(1 + 240 \sum_{n=1}^{\infty} \sigma_3(n) e^{-2\pi n}\right)$$

which is only possible if $c = 0$. Hence $f = 0$. ■

Now we get a complete picture of the dimension of M_k for all k.

Theorem 4.5. *For even $k > 0$ the dimension of M_k is*

$$\dim M_k = \begin{cases} \lfloor \frac{k}{12} \rfloor + 1, & \text{if } k \not\equiv 2 \bmod 12 \\ \lfloor \frac{k}{12} \rfloor, & \text{if } k \equiv 2 \bmod 12 \end{cases}$$

Proof. Define a $\mathbb{C}$-linear map $\mathbb{C} \oplus M_{k-12} \to M_k$ by $(c, f) \mapsto cE_k + f\Delta_r$. This is surjective by 4.3. To show it is injective suppose $cE_k + f\Delta_r$ is 0. Then the constant term in its q expansion must be 0, because the constant term in the q expansion of Δ_r is 0 and the constant term of E_k is 1 it follows that if the constant term of $cE_k + f\Delta_r$ is 0 c is 0. But now its just $f\Delta_r$ which vanishes iff either f or Δ_r vanishes. But $\Delta_r \neq 0$ so $f = 0$. Hence the kernel of our linear map is 0.

Therefore $M_k \cong M_{k-12} \oplus \mathbb{C}$ so we get a recurrence

$$\dim M_k = 1 + \dim M_{k-12}$$

the claim then follows easily from induction on this recurrence with our base case given in 4.4. ■

Remark 4.6. It is notable to include that because the condition that f be a cusp form is a condition on exactly one of its Fourier coefficients (namely that $a_0 = 0$) the dimension of S_k is equal to $\dim M_k - 1$, an immediate corrolary to this is that M_k decomposes as

$$\langle E_k \rangle \oplus S_k$$

Now that we have a formula for dimension we can show a truly miraculous structure theorem.

Theorem 4.7.

$$\mathbb{C}[E_4, E_6] \cong \bigoplus M_k$$

Proof. We show that the set $\{E_4^a E_6^b : 4a + 6b = k\}$ forms a basis for M_k. First we prove the number of solutions to $4a + 6b$ is the same as the dimension of M_k. Assume $k \geq 12$, let us look at solutions (a, b) with $b \geq 2$, then $(a, b-2)$ is a solution to $4a + 6b = k - 12$, the map $(a, b) \mapsto (a, b-2)$ is injective and in fact bijective because if (a', b') solves $4a + 6b = k - 12$ then certainly $(a', b' + 2)$ solves $4a + 6b = k$. Now if $b < 2$ then either $b = 1$ or $b = 0$. If $k \equiv 2 \bmod 4$ then $b \neq 0$ and as $k - 6 \equiv 0 \bmod 4$ there is exactly one solution $(a, 1)$. If $k \equiv 0 \bmod 4$ then $b \neq 0$ and for some a $k = 4a$ so there is one solution $(a, 0)$. Hence there is one more solution to $4a + 6b = k$ than there is to $4a + 6b = k - 12$. If N_k represents the number of solutions to this equation for k we obtain $N_k = N_{k-12} + 1$. Via directly checking we also have that for $k \leq 10$ $N_k = \dim M_k$, hence N_k has the same recurrence and base cases as $\dim M_k$ so they are equal.

Now we show $\{E_4^a E_6^b : 4a + 6b = k\}$ is linearly independent. Suppose they are linearly dependent, then there is some

$$\sum_{4a+6b=k} c_{a,b} E_4^a E_6^b = 0$$

With not all $c_{a,b} = 0$, we induct on k, the base case is trivial as for $k \leq 10$ there is only one element of $\{E_4^a E_6^b : 4a + 6b = k\}$. Now note that

$$E_6(i) = E_6(\frac{-1}{i}) = i^6 E_6(i) = -E_6(i)$$

so $E_6(i) = 0$. However $E_4(i) \neq 0$. Therefore by plugging i into

$$\sum_{4a+6b=k} c_{a,b} E_4^a E_6^b$$

We find that any pure E_4 term (one without a factor of E_6) must have $c_{a,0} = 0$. Hence every non-zero terms has an E_6 factor, factoring this out and cancelling gives a sum

$$\sum_{4a+6b=k-6} c_{a,b} E_4^a E_6^b = 0$$

which contradicts the induction hypothesis. ∎

Remark 4.8. We now know that we can describe any modular form as a sum of these $E_4^a E_6^b$ terms. Whats awesome is that this essentially trivializes finding q-expansions of modular forms. Let us look at

$$\Delta_r = q \prod_{n=1}^{\infty} (1 - q^n)^{24}$$

Then as it is a cusp form the first term of its q-expansion is $a_0 = 0$ and clearly from its definition $a_1 = 1$. But, because it is weight 12, we know it is a sum of $aE_4^3 + bE_6^2$. The constant terms of E_4^3 and E_6^2 are both 1 and so we must have

$$\Delta_r = c(E_4^3 - E_6^2)$$

Now

$$E_6^2 = \left(1 - 504 \sum_{n=1}^{\infty} \sigma_5(n) q^n\right)^2 = 1 + \sum_{n=1}^{\infty} \left(-1008\sigma_5(n) + \sum_{k=1}^{\infty} 504^2 \sigma_5(k)\sigma_5(n-k)\right) q^n$$

And

$$E_4^3 = \left(1 + 240 \sum_{n=1}^{\infty} \sigma_3(n) q^n\right)^3 = 1 + \sum_{n=1}^{\infty} \left(720\sigma_3(n) + 3(240^2) \sum_{k=1}^{n-1} \sigma_3(k)\sigma_3(n-k) + 3(240)^3 \sum_{m+j+k=n} \sigma_3(k)\sigma_3\right.$$

So the q term in each of these q expansions is

$$[q]E_6^2 = -1008 \qquad [q]E_4^3 = 720$$

so the q coefficient of $E_4^3 - E_6^2$ is 1728. So because we need the q term in Δ_r to be 1 we therefore have

$$\Delta_r = \frac{E_4^3 - E_6^2}{1728}$$

The coefficients of the q expansion of Δ_r are denoted $\tau(n)$ and known as the Ramanujan tau function. Using what we have just found we can calculate the first few terms of the tau function Now the very astute reader might have noticed that $\tau(6) = \tau(2)\tau(3)$ and

n	1	2	3	4	5	6	7	8	9	10
$\tau(n)$	1	-24	252	-1472	4830	-6048	-16744	84480	-113643	-115920

$\tau(10) = \tau(5)\tau(2)$ and indeed if we were to continue on this table forever we would find that for any $\gcd(n, m) = 1$ $\tau(nm) = \tau(n)\tau(m)$. To do so we will introduce something called Hecke operators.

5. Hecke Operators

The goal of this section will be to build a sequence of commuting operators T_n on the space M_k which have a number of nice properties, then transfer these properties to certain modular forms (called eigenforms) including Δ_r and be able to conclude things about their coefficients.

First though, recall how in the first section we showed an equivalence between modular forms of weight k and functions F on the space of complex lattices such that for any complex number λ we have

$$F(\lambda\Lambda) = \lambda^{-k}F(\Lambda)$$

with this idea in mind we come up with the following fairly natural definition.

Definition 5.1. For any $n \in \mathbb{N}$ define T_n by its action on F

$$T_nF(\Lambda) = n^{k-1}\sum_{[\Lambda:\Lambda']=n} F(\Lambda')$$

Since there are only finitely many Λ' with $[\Lambda : \Lambda'] = n$ this is well defined.

This descends to modular forms f via

$$T_nf(z) = n^{k-1}\sum_{\begin{bmatrix} a & b \\ c & d\end{bmatrix}\in \mathrm{SL}_2(\mathbb{Z})n\mathcal{M}_n} (cz+d)^{-k}f(\frac{az+b}{cz+d})$$

where $\mathrm{SL}_2(\mathbb{Z})n\mathcal{M}_n$ is the set of orbits of $\mathcal{M}_n$, the space of integer matrices with determinant n, under left multiplication by $\mathrm{SL}_2(\mathbb{Z})$. We call T_n the nth Hecke operator.

A straightforward but somewhat lengthy check shows that these definitions do indeed coincide. We omit details but for readers who wish to check it themselves the key point is that index n sublattices correspond to cosets $\mathrm{SL}_2(\mathbb{Z})n\mathcal{M}_n$.

Theorem 5.2. *If $f(z)$ is a modular form of weight k with Fourier expansion $\sum_{m=0}^{\infty} a_mq^m$ then the Fourier expansion of T_nf is given by*

$$\sum_{m=0}^{\infty} q^m \sum_{d|n,m} d^{k-1}a_{nm/d^2}$$

In particular T_n sends modular forms of weight k to modular forms of weight k and cusp forms to cusp forms.

Proof. We begin by choosing suitable representatives for the cosets in $\mathrm{SL}_2(\mathbb{Z})n\mathcal{M}_n$. If $\mu = \begin{bmatrix} a & b \\ c & d\end{bmatrix}$ has determinant n and has $c \neq 0$ choose $\gamma = \begin{bmatrix} a' & b' \\ c' & d'\end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ such that $\frac{a'}{c'} = \frac{a}{c}$. Then the lower left entry of $\gamma^{-1}\mu$ is 0, hence we may choose coset representatives $\begin{bmatrix} a & b \\ 0 & d\end{bmatrix}$. Now if $\mu' = \begin{bmatrix} e & f \\ 0 & g\end{bmatrix}$ is another such representative we must have

$$\gamma\mu = \mu'$$

, but as both have 0 in the lower left entry we must have $\gamma = \pm \begin{bmatrix} 1 & l \\ 0 & 1 \end{bmatrix}$ in which case $\mu' = \begin{bmatrix} a & b+ds \\ 0 & d \end{bmatrix}$ So our choice μ is unique if we require $a, d > 0$ and $0 \leq b < d$. Therefore we have

$$T_n f(z) = n^{k-1} \sum_{d|n} \sum_{b=0}^{d-1} d^{-k} f(\frac{nz+db}{d^2})$$

Now

$$f(\frac{nz+db}{d^2}) = \sum_{m=0}^{\infty} a_m e^{2\pi imnz/d^2} e^{2\pi imb/d}$$

and

$$\sum_{b=0}^{d-1} e^{2\pi imb/d} = \begin{cases} d, & d \mid m \\ 0, & d \nmid m \end{cases}$$

So we only care about those d which divide m. Hence the sum becomes

$$n^{k-1} \sum_{m=0}^{n} \sum_{d|n,m} d^{1-k} a_m \exp(2\pi imnz/d^2)$$

which we may rearrange via first substituting $m/d \to m$ and then $\frac{n}{d} \to d$ to obtain

$$\sum_{m=0}^{n} q^m \sum_{d|n,m} d^{k-1} a_{nm/d^2}$$

As desired. ∎

Corollary 5.3. *The Hecke operators T_m, T_n (of weight k) satisfy*

$$T_n T_m = \sum_{d|n,m} d^{k-1} T_{mn/d^2}$$

And in particular hecke operators commute and if n, m are coprime $T_n T_m = T_{nm}$.

Proof. Take any weight k form $f(z) = \sum_{r=0}^{\infty} a_r q^r$. The action on the t-th coefficient from $T_n T_m$ is then

$$\sum_{d_1|n,t} \sum_{d_2|nt/d_1^2,m} (d_1 d_2)^{k-1} a_{nmt/(d_1 d_2)^2} = \sum_{d_2|nt/(d_1^2),d_2|m,d_1|n,d_1|t} (d_1 d_2)^{k-1} a_{nmt/(d_1 d_2)^2}$$

And the action of T_{nm/d^2}

$$\sum_{d_3|mn/d^2,t} d_3^{k-1} a_{nmt/d^2 d_3^2}$$

Now substitute $d = \gcd(d_1, d_2)$ and $e = \frac{d_1 d_2}{d}$. Then $d \mid n, m$, $e \mid \frac{mn}{d^2}$, $e \mid t$. Hence the first sum becomes

$$\sum_{d|m,n} d^{k-1} \sum_{e|(mn/d^2),s} e^{k-1} a_{\frac{mnt}{d^2 e^2}} = \sum_{d|n.m} d^{k-1} ([q] T_{mn/d^2} f)$$

hence

$$T_m T_n = \sum_{d|n,m} d^{k-1} T_{mn/d^2}$$

∎

Now that we have these few facts we are able to talk about a certain nice class of modular forms.

Definition 5.4. A modular form f is an eigenform if for each T_n there is a scalar λ_n with

$$T_n f = \lambda_n f$$

. We call a form f a normalized eigenform if it has $a_1 = 1$. We can write the condition that f be a Hecke eigenform more specifically as

$$\lambda_n a_m = \sum_{d|n,m} d^{k-1} a_{nm/d^2}$$

Which in particular gives the extremely useful fact that $\lambda_n a_1 = a_n$.

Remark 5.5. You might think, with how complicated Hecke operators seem to be, that eigenforms are uncommon. This is far from the truth, because those $k \leq 10$ have M_k one dimensional and $T_n : M_k \to M_k$ every weight $k \leq 10$ modular form is an eigenform. We can further extend this. Hecke operators preserve cusp forms and the space of cusp forms has dimension one less than that of M_k, this tells us that all cusp forms of weight $k = 12, 16, 18, 20, , 22, 26$ are eigenforms. In particular this includes Δ_r which is a normalized eigenform.

Normalized eigenforms are extremely nice to work with and we are able to give a recurrence on the Fourier coefficients of a normalized eigenform.

Proposition 5.6. *If f is a normalized eigenform then its Fourier coefficients a_n satisfy.*

$$a_n a_m = \sum_{d|n,m} d^{k-1} a_{nm/d^2}$$

And so in particular its coefficients are multiplicative (ie $a_n a_m = a_{nm}$ whenever n, m are coprime) and for a prime power p^r they satisfy (via setting $n = p^r, m = p$)

$$a(p^{r+1}) = a(p)a(p^r) - p^{k-1}a(p^{r-1})$$

Proof. This is immediate from the fact that $a_1 = 1$ $\lambda_n a_1 = a_n$ and the characterization of eigenforms by

$$\lambda_n a_m = \sum_{d|n,m} d^{k-1} a_{nm/d^2}$$

■

6. CONCLUSION

We have developed the basic theory of modular forms, beginning with their definition and proceeding through their Fourier expansions and fundamental examples. Using the discriminant modular form, we obtained a recursive description of the spaces M_k, leading to the structure theorem that the graded algebra of modular forms is generated by E_4 and E_6. We then introduced Hecke operators and showed how they constrain the Fourier coefficients of modular forms, particularly in the case of eigenforms.

Several important aspects of the theory have been omitted. In particular, we have not treated modular forms on congruence subgroups, nor their connections to elliptic curves, L-functions, or representation theory. We have also avoided a systematic study of Hecke algebras and their spectral properties. These topics lie beyond the scope of this paper, but the results presented here provide a foundation for their study.