Gross–Zagier reading seminar: Modular Forms and Hecke Operators

Lecture 4 • Charlotte Chan • September 30, 2014

1. Introduction

Our goal is to give an introduction to modular forms and to discuss the role of Hecke operators in the theory. We will focus on the parts of the theory that we will need in studying [GZ86]. For more details, one can refer to [DI95].

Recall the following topics from Jeff's talks:

- The Atkin-Lehner involution W_k is an involution on the modular curve $Y_0(N) = \mathfrak{h}/\Gamma_0(N)$ and permute the cusps.
- Eisenstein series...
- A *Heegner point* is a point in the modular curve $Y_0(N) = \mathfrak{h}/\Gamma_0(N)$ corresponding to N-isogenous elliptic curves E, E' with CM by the same order $\mathcal{O}$.

In this talk, we will begin by introducing modular forms. We will then define the Hecke algebra[1] together with its action on the following objects: the space of weight-k modular forms with respect to $\Gamma_0(N)$ and (divisors of) the modular curve $Y_0(N)$. The moduli space interpretation of $Y_0(N)$ allows us then interpret the latter action as a modular correspondence—the mth Hecke operator $T(m)$ will take a pair (E, C) consisting of an elliptic curve E and a cyclic order-N subgroup $C \subset E$ to the average over pairs $(E/D, C + D/D)$ where $D \subset E$ is a cyclic order-m subgroup with $C \cap D = \varnothing$. In particular, we will see that $T(m)$ acts on the set of divisors of $Y_0(N)$ supported on the Heegner points.

Remark. Later on this semester, we will need to prove that for a Heegner point c, the power series

$$\sum_{n\geq 1} \langle c, T_n c\rangle \, q^n$$

is a modular form. We will see that the Hecke algebra $\mathbb{T}$ acts on the space of weight-2 cusp forms $\mathcal{S}_2(\Gamma)$ for $\Gamma = \Gamma_0(N)$ and on the Jacobian $\mathrm{Jac}(X_0(N))$. Thus we have maps $\varphi_1 \colon \mathbb{T} \to \mathrm{End}(\mathcal{S}_2(\Gamma))$ and $\varphi_2 \colon \mathbb{T} \to \mathrm{End}(\mathrm{Jac}(X_0(N)))$. The map $T_n \mapsto \langle c, T_n c\rangle$ gives a ring homomorphism $\varphi_2(\mathbb{T}) \to \mathbb{C}$. It turns out that this map also defines a ring homomorphism $\varphi_1(\mathbb{T}) \to \mathbb{C}$. By the discussion in Section 4, we see that this then implies that $\sum_{n\geq 1}\langle c, T_n c\rangle \, q^n$ is indeed a modular form. ◊

2. Modular forms

The subgroup $\mathrm{GL}_2^+(\mathbb{R}) \subset \mathrm{GL}_2(\mathbb{R})$ consisting of positive-determinant matrices acts on the upper-half plane $\mathfrak{h}$ via Möbius transformations. That is, for $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}_2^+(\mathbb{R})$ and $z \in \mathfrak{h}$,

$$\gamma \cdot z := \frac{az + b}{cz + d}.$$

[1] The Hecke algebra $\mathbb{T}_N$ that we will define acts on the space of weight-k cusp forms $\mathcal{S}_k(\Gamma)$, giving a map $\mathbb{T}_N \to \mathrm{End}(\mathcal{S}_k(\Gamma))$. Sometimes (though not in this talk) we will also call the image of $\mathbb{T}_N$ the Hecke algebra.

For any $k \in \mathbb{Z}$, we can define a $\mathrm{GL}_2^+(\mathbb{R})$-action on the space of functions $f\colon \mathfrak{h} \to \mathbb{C}$. This action is given by $\gamma \cdot f := f|_{[\gamma]_k}$, where

$$f|_{[\gamma]_k}(z) := \det(\gamma)^{k-1}(cz+d)^{-k} f(\gamma \cdot z).$$

What we are interested in is the study of the action of certain subgroups $\Gamma \subset \mathrm{GL}_2^+(\mathbb{R})$ on certain functions $f\colon \mathfrak{h} \to \mathbb{C}$.

2.1. **Modular forms and cusp forms.** Consider the following subgroups of $\mathrm{SL}_2(\mathbb{Z})$:

$$\Gamma(N) = \{\gamma \in \mathrm{SL}_2(\mathbb{Z}) : \gamma \equiv \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \mod N\}$$
$$\Gamma_1(N) = \{\gamma \in \mathrm{SL}_2(\mathbb{Z}) : \gamma \equiv \left(\begin{smallmatrix} 1 & * \\ 0 & 1 \end{smallmatrix}\right) \mod N\}$$
$$\Gamma_0(N) = \{\gamma \in \mathrm{SL}_2(\mathbb{Z}) : \gamma \equiv \left(\begin{smallmatrix} * & * \\ 0 & * \end{smallmatrix}\right) \mod N\}$$

A subgroup $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ is a *congruence subgroup* if there exists some $N \in \mathbb{N}$ for which Γ contains $\Gamma(N)$. For example, for any $N \in \mathbb{N}$, the subgroups $\Gamma_1(N)$ and $\Gamma_0(N)$ are congruence subgroups.

Definition 1. A *modular form* of weight $k \in \mathbb{Z}_{\geq 0}$ with respect to a congruence subgroup Γ is a function $f\colon \mathfrak{h} \to \mathbb{C}$ satisfying

(i) f is holomorphic on $\mathfrak{h}$
(ii) $f = f|_{[\gamma]_k}$ for all $\gamma \in \Gamma$
(iii) f is holomorphic at the cusps

The space of modular forms of weight k with respect to Γ is denoted $\mathcal{M}_k(\Gamma)$.

Definition 2. A *cusp form* of weight $k \in \mathbb{Z}_{\geq 0}$ with respect to a congruence subgroup Γ is function $f \in \mathcal{M}_k(\Gamma)$ that vanishes at the cusps. The subspace of cusp forms in $\mathcal{M}_k(\Gamma)$ is denoted $\mathcal{S}_k(\Gamma)$.

Remark. We elaborate on Definition 1(iii) and Definition 2. Fix a non-negative integer k and a congruence subgroup Γ. Suppose $f\colon \mathfrak{h} \to \mathbb{C}$ satisfies $f = f|_{[\gamma]_k}$ for all $\gamma \in \Gamma$. Then for some $h > 0$, we have $f(z) = f(z+h)$ for all $z \in \mathfrak{h}$. (This holds since there exists some $h > 0$ such that $\left(\begin{smallmatrix} 1 & h \\ 0 & 1 \end{smallmatrix}\right) \in \Gamma$.) Thus f has a Fourier expansion at ∞ given by

$$f(z) = \sum_{n=-\infty}^{\infty} a_n q_h^n, \qquad \text{where } q_h = e^{2\pi i z/h}.$$

We say that f is *holomorphic* (resp. *vanishes*) *at* ∞ if $a_n = 0$ for $n < 0$.

Since $\alpha \cdot f := f|_{[\alpha]_k}$ defines an action, then for any $\alpha \in \mathrm{SL}_2(\mathbb{Z})$ and any $\gamma \in \alpha^{-1}\Gamma\alpha$, we have $f|_{[\alpha]_k}|_{[\gamma]_k} = f|_{[\alpha]_k}$. Thus $f|_{[\alpha]_k}$ also has a Fourier expansion at ∞ and we can talk about $f|_{[\alpha]_k}$ being holomorphic (resp. vanishing) at ∞.

We say f is *holomorphic* (resp. *vanishes*) *at the cusps* if $f|_{[\alpha]_k}$ is holomorphic at ∞ for all $\alpha \in \mathrm{SL}_2(\mathbb{Z})$. ◊

Remark. Suppose $k = 0$. If $f \in \mathcal{M}_0(\Gamma)$, then f can be viewed as a holomorphic function from a compact space to $\mathbb{C}$. If f is nonconstant, the image of f must be a compact open subset of $\mathbb{C}$. This is a contradiction and we see that $\mathcal{M}_0(\Gamma) = \mathbb{C}$ and $\mathcal{S}_0(\Gamma) = \{0\}$. ◊

Proposition 3. *The space $\mathcal{M}_k(\Gamma)$ is finite-dimensional.*

From this point forward, we will set $\Gamma := \Gamma_0(N)$ and N will always refer to the level of Γ. We will also set $Y := Y_0(N)$.

3. The Hecke algebra and two actions

Consider

$$\Delta := \Delta_0(N) = \{\alpha \in M_2(\mathbb{Z}) : \det(\alpha) > 0,\, \alpha \equiv \left(\begin{smallmatrix} * & * \\ 0 & * \end{smallmatrix}\right) \mod N,\, (a(\alpha), N) = 1\}$$

Now let $R(\Gamma, \Delta)$ denote the $\mathbb{Z}$-module generated by the double cosets $\Gamma\alpha\Gamma$ for $\alpha \in \Delta$. (If $\alpha, \alpha' \in \Delta$ are such that $\Gamma\alpha\Gamma = \Gamma\alpha'\Gamma$, we consider these to be the same element in $R(\Gamma, \Delta)$.) One can endow $R(\Gamma, \Delta)$ with a multiplication structure by studying the coset decompositions $\Gamma\alpha\Gamma = \sqcup_i \Gamma\alpha_i$. (See, for example, [DI95].) Equipped with this multiplication law, $R(\Gamma, \Delta)$ is an associative, commutative ring with identity (namely $\Gamma = \Gamma \cdot 1 \cdot \Gamma$).

3.1. Two actions.

The Hecke ring $R(\Gamma, \Delta)$ acts on $\mathcal{M}_k(\Gamma)$ and $\mathrm{Pic}(Y)$ via the following rule. For $\alpha \in \Delta$, write $\Gamma\alpha\Gamma = \sqcup_i \Gamma\alpha_i$ and define

$$f|_{[\Gamma\alpha\Gamma]_k} = \sum_i f|_{[\alpha_i]_k} \qquad \text{for } f \in \mathcal{M}_k(\Gamma),$$
$$(\Gamma\alpha\Gamma) \cdot x = \sum_i \alpha_i \cdot z \qquad \text{for } z \in \mathfrak{h}/\Gamma.$$

This defines an action of $R(\Gamma, \Delta)$ on $\mathcal{M}_k(\Gamma)$ and $\mathrm{Pic}(Y)$. Notice that $\mathcal{S}_k(\Gamma)$ is stable under this action. Later, we will see that the set of divisors supported on Heegner points is stable under this Hecke action. These actions can be made very explicit and will ultimately allow us to understand:

(i) the relationship between the Fourier coefficients of f and eigenvalues for certain elements of $R(\Gamma, \Delta)$ called Hecke operators
(ii) Hecke operators on N-isogenous elliptic curves as averages over isogenies

We will discuss (i) in Section 4 and (ii) in Section 5.

Let $\Delta^n = \{\alpha \in \Delta : \det \alpha = n\}$ and define the *Hecke operator* $T_m := T(m)$ as

$$T_m := T(m) := \sum_{\substack{\alpha \in \Delta^m, \\ \Gamma\alpha\Gamma \text{ distinct}}} \Gamma\alpha\Gamma.$$

It turns out that $R(\Gamma, \Delta)$ is the polynomial ring $\mathbb{T}_N$ over $\mathbb{Z}$ generated by $T(p) = \Gamma \left(\begin{smallmatrix} 1 & 0 \\ 0 & p \end{smallmatrix}\right) \Gamma$ and $T(p, p) := \Gamma \begin{pmatrix} p & 0 \\ 0 & p \end{pmatrix} \Gamma$ for all primes p. We will also be interested in the subring $\mathbb{T}^{(N)} \subset \mathbb{T}_N$ generated by $T(p)$ and $T(p, p)$ for all primes p not dividing N.

The following lemma is very useful in explicitly studying Hecke operators.

Lemma 4 (Shimura). *For every $m \in \mathbb{N}$, we have*

$$\Delta^m = \bigsqcup_a \bigsqcup_b \Gamma \left(\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix} \right)$$

where the disjoint union is over $a > 0$ coprime to N with $ad = m$ and over $b \pmod{d}$.

4. Hecke operators on modular forms

4.1. **Fourier coefficients and Hecke eigenvalues.** By Lemma 4, for $f \in \mathcal{M}_k(\Gamma)$,

$$f|_{T(m)_k}(z) := n^{k-1} \sum_a \sum_{b=0}^{d-1} d^{-k} f\left(\frac{az+b}{d}\right) \qquad (a > 0,\, ad = m). \tag{1}$$

Now let $\sum_{n\geq 0} c_n q^n$ be the q-expansion of $f \in \mathcal{M}_k(\Gamma)$ and let $\sum_{n\geq 0} d_n q^n$ be the q-expansion of $f|_{T(m)_k}$. We wish to compute d_n using Equation (1).

Let $q = e^{2\pi i z/h}$. Then

$$\begin{aligned} f|_{T(m)_k}(z) &= m^{k-1} \sum_a \sum_{b=0}^{d-1} d^{-k} f\left(\frac{az+b}{d}\right) \\ &= m^{k-1} \sum_a \left(\frac{m}{a}\right)^{-k} \sum_{b=0}^{m/a-1} \sum_{n\geq 0} c_n e^{\frac{2\pi iab}{mh}} q^{na^2/m}. \end{aligned}$$

Collecting coefficients of q^n, it follows that

$$\begin{aligned} d_n &= m^{k-1} \sum_{a|(m,n)} \left(\frac{m}{a}\right)^{-k} \sum_{b=0}^{m/a-1} c_{nm/a^2} e^{\frac{2\pi iab}{mh}} \\ &= m^{k-1} \sum_{a|(m,n)} \left(\frac{m}{a}\right)^{-k+1} c_{nm/a^2} \\ &= \sum_{a|(m,n)} a^{k-1} c_{nm/a^2}. \end{aligned} \tag{2}$$

Remark. The computation of Equation (2) was very elementary, but it already tells us something about the relationship between Hecke operators and the Fourier coefficients of modular forms. Suppose that $f \in \mathcal{M}_k(\Gamma)$ is an eigenform $\{T(m)\}_{m\in S}$ for some subset $S \subset \mathbb{N}$. If λ_m is the $T(m)$-eigenvalue for f, and the q-expansion of f is $\sum_{n\geq 0} c_n q^n$, then Equation (2) implies

$$\lambda_m c_1 = d_1 = \sum_{a|(m,1)} a^{k-1} c_{m/a^2} = c_m.$$

In particular, if $f \in \mathcal{M}_k(\Gamma)$ is an eigenform for $\{T(m)\}_{m\in\mathbb{N}}$. Then the $T(m)$-eigenvalues of f determine f up to scaling. ◊

4.2. **Old and new subspaces.** For d, M such that dM divides N, the action of $\iota_d = \left(\begin{smallmatrix} d & 0 \\ 0 & 1 \end{smallmatrix}\right)$ defines an injective map

$$\iota^*_{d,M,N} \colon \mathcal{S}_k(\Gamma) \to \mathcal{S}_k(\Gamma), \qquad\qquad f \mapsto f|_{[\iota_d]_k}.$$

Explicitly, $f|_{[\iota_d]_k}(z) = d^{k-1} f(dz)$. (Note that the image of $\iota^*_{d,M,N}$ is in $\mathcal{S}_k(\Gamma)$ since $\iota_d^{-1}\Gamma\iota_d$ contains Γ.)

We define the *old subspace* $\mathcal{S}_k(\Gamma)^{\text{old}}$ to be the span of all possible images of $\mathcal{S}_k(\Gamma)^{\text{old}}$ via $\iota^*_{d,M,N}$. We then define the *new subspace* $\mathcal{S}_k(\Gamma)^{\text{new}}$ to be the orthogonal complement of $\mathcal{S}_k(\Gamma)^{\text{old}}$ with respect to the Petersson inner product

$$\langle -, - \rangle \colon \mathcal{S}_k(\Gamma) \times \mathcal{S}_k(\Gamma) \to \mathbb{C}$$

defined as

$$\langle f, g \rangle := \frac{1}{[\mathrm{PSL}_2(\mathbb{Z}) : \Gamma/\{\pm 1\}]} \int_D f(z)\overline{g(z)} y^k \frac{dx\,dy}{y^2},$$

where D is a fundamental domain for $\Gamma = \Gamma_0(N)$.

Note that $\iota^*_{d,M,N}$ commutes with the action of $T(p)$ if $p \nmid N$. In particular, this implies that if $f \in \mathcal{S}_k(\Gamma)$ is a $\mathbb{T}^{(N)}$-eigenform, then $\iota_{d,M,N}(f) \in \mathcal{S}_k(\Gamma)$ is again a $\mathbb{T}^{(N)}$-eigenform with the same $\mathbb{T}^{(N)}$-eigenvalues.

Theorem 5 (Multiplicity One)**.** *Let $f, g \in \mathcal{S}_k(\Gamma)$ be $\mathbb{T}^{(ND)}$-eigenforms with the same eigencharacters. If $f \in \mathcal{S}_k(\Gamma)^{new}$ is normalized, then g is a scalar multiple of g.*

Fact 6. *$\mathcal{S}_k(\Gamma)^{new}$ has a basis consisting of $\mathbb{T}^{(N)}$-eigenforms.*

Theorem 5 together with Fact 6 imply

Corollary 7. *The subspace $\mathcal{S}_k(\Gamma)$ is the orthogonal sum of $\mathbb{T}^{(ND)}$-eigenspaces. Furthermore, a $\mathbb{T}^{(ND)}$-eigenspace V is in the new subspace if and only if $\dim V = 1$.*

Note that this corollary implies that $\mathcal{S}_k(\Gamma)^{\text{new}}$ is stable under the action of $\mathbb{T}_N$ and that each new $\mathbb{T}^{(ND)}$-eigenspace is in fact a $\mathbb{T}_N$-eigenspace. Any form in this space is called a *newform* and the one whose first Fourier coefficient is 1 is the normalized new form.

5. Hecke operators as modular correspondences

By Lemma 4, if $(p, N) = 1$, then for $\tau \in \mathfrak{h}$,

$$T_p(\Gamma\tau) = \sum_\gamma \Gamma\gamma(\tau),$$

where the sum ranges over $\gamma \in \left\{ \left(\begin{smallmatrix} 1 & 0 \\ 0 & p \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 & 1 \\ 0 & p \end{smallmatrix}\right), \ldots, \left(\begin{smallmatrix} 1 & p-1 \\ 0 & p \end{smallmatrix}\right), \left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right) \right\}$. Recall that the action of γ on $\tau \in \mathfrak{h}$ is given by Möbius transformations. Equivalently, for $x \in Y$,

$$T_p(x) = \sum_\gamma \gamma \cdot x,$$

where the sum ranges over γ in the same set as before.

Remark. Here is a small example that foreshadows the next discussion. For $\tau \in \mathfrak{h}$, we have $\left(\begin{smallmatrix} 1 & 1 \\ 0 & p \end{smallmatrix}\right) \cdot \tau = (\tau+1)/p$, which corresponds to the elliptic curve $\mathbb{C}/\langle 1, (\tau+1)/p \rangle = (\mathbb{C}/\langle 1, \tau \rangle)/D$, where D is the order-p subgroup of $\mathbb{C}/\langle 1, \tau \rangle$ generated by $(\tau+1)/p$. Viewing τ in the quotient $\mathfrak{h}/\Gamma_0(N) = Y$, the moduli interpretation of Y allows us to associate to τ the pair (E, C), where $E = \mathbb{C}/\langle 1, \tau \rangle$ and $C \subset E$ is a cyclic subgroup of order N. Then $\left(\begin{smallmatrix} 1 & 1 \\ 0 & p \end{smallmatrix}\right) \cdot \tau \in Y$ corresponds to the pair $(E/D, (C+D)/D)$. $\Diamond$

Using the moduli space interpretation of $Y = Y_0(N)$, we know that a point $x \in Y$ corresponds to an N-isogeny $E \to E'$ of elliptic curves. Equivalently, $x \in Y$ corresponds to a pair (E, C) where $C \subset E$ is a cyclic group of order N. Then

$$T_p(E, C) = \sum_D (E/D, (C+D)/D)$$

where the sum ranges over order-p subgroups $D \subset E$.

In general, for any $n \in \mathbb{N}$,

$$T_n(E, C) = \sum_D (E/D, (C+D)/D),$$

where the sum ranges over cyclic order-n subgroups $D \subset E$ such that $C \cap D = 0$.

Remark. If (E, C) is a Heegner point, then E and E/C have CM by the same order $\mathcal{O}$. Then E/D and $E/(C+D)$ have CM by the same order. To show this, one can use the discriminant criterion for a Heegner point that Jeff discussed in the previous talk. $\Diamond$

Remark. The correspondence given by T_p can also be described in the following way. A point in $Y_0(Np)$ is a pair (E, C) where $C \subset E$ is a cyclic subgroup of order Np. Then we can define maps

$$\begin{aligned} \alpha &\colon Y_0(Np) \to Y_0(N) & \qquad (E, C) &\mapsto (E, C_N), \\ \beta &\colon Y_0(Np) \to Y_0(N) & \qquad (E, C) &\mapsto (E/C_p, C/C_p), \end{aligned}$$

where $C_p, C_N \subset C$ is the subgroup of order p, N, respectively. This gives a diagram

$$\begin{array}{ccccc} & & Y_0(Np) & & \\ & \swarrow^{\alpha} & & \searrow^{\beta} & \\ Y_0(N) & & & & Y_0(N) \end{array}$$

Then the Hecke operator T_p on $\mathrm{Div}(Y_0(N))$ is the composition $T_p = \alpha \circ {}^t\beta$, where ${}^t\beta$ sends a point $x \in Y_0(N)$ to the formal sum over its preimages in $Y_0(Np)$. $\Diamond$

References

[DI95] Diamond, F. and Im, J. *Modular forms and modular curves.* Canadian Mathematical Society, Conference Proceedings, Volume 17, 1995.

[G84] Gross, B. *Heegner points on $X_0(N)$.* In: Modular Forms (ed. R.A. Rankin). 1984.

[GZ86] Gross, B. and Zagier, D. *Heegner points and derivatives of L-series.* Inventiones, 1986.