EICHLER-SHIMURA THEORY AND THE EXISTENCE OF RANK 0 QUOTIENTS OF MODULAR JACOBIANS

FRANK THORNE

Abstract. These are rough notes for a talk to be given in the Wisconsin student number theory seminar. The first is a broad overview of Eichler-Shimura theory, giving some explanation of how one can associate elliptic curves to modular forms. The second describes how an analytic argument can be used to establish the existence of certain rank 0 quotients of modular Jacobians, and how this fact fits into an argument of Mazur.

Disclaimer: While the author has attempted to be clear and correct, this paper was produced by a nonexpert in a short amount of time. The reader who would like to learn this material in any serious fashion is advised to stop reading immediately and buy Knapp's book instead.

1. Introduction

A **modular form** is a certain type of analytic function which obeys particular transformation properties on the upper half plane. An **elliptic curve** is a nonsingular projective curve of genus 1, and it turns out they can be described in terms of simple Weierstrass equations.

It is somewhat surprising that these objects have everything to do with each other. In particular, the following is true:

Theorem 1.1. *There is a bijective correspondence between newforms of weight 2 with integer Fourier coefficients, and isogeny classes of elliptic curves defined over $\mathbb{Q}$.*

This is amazing! A lot more is true: elliptic curves of conductor N correspond to cusp forms in $\Gamma_0(N)$, and the associated L-functions match up under this correspondence.

Perhaps the most famous direction of this correspondence is the famous

Theorem 1.2. *("Taniyama-Shimura Conjecture") Given an elliptic curve $E/\mathbb{Q}$, one can associate to it a modular form f in a canonical way. (In particular, the L-functions have to match...)*

The proof of this is ridiculously difficult and involves a great deal of technical machinery (as those of you in Nigel's seminar know...) One obtains FLT as a corollary. One might ask, does the correspondence go the other way? This is also difficult, but this direction was known previously.

Theorem 1.3 (Eichler-Shimura)**.** *Given a newform f of weight 2, with integer coefficients, one can associate a unique elliptic curve $E/\mathbb{Q}$ so that the L-functions will match.*

The first portion of this talk will attempt to give an overview of Eichler-Shimura theory, following the book by Knapp.

2. Example: $X_0(11)$

Everyone's favorite example of an elliptic curve is $X_0(11)$. (Well, everyone except Tunnell, Koblitz, and Heath-Brown.) This curve may be given by the equation

$$y^2 + y = x^3 - x^2 - 10x - 20.$$

2000 *Mathematics Subject Classification.* 11N25, 11N36.

Now we make the usual definitions

$$\Gamma_0(N) = \{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(Z) | \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} (mod N)\}.$$

and

$$\mathbb{H} = \{x \in \mathbb{C} : Im(x) > 0.\}$$

$\Gamma_0(N)$ acts on $\mathbb{H}$ by fractional linear transformations, and the resulting quotient space is called $Y_0(N)$. We compactify it by adding cusps, and we obtain a Riemann surface. This compactification is referred to as $X_0(N)$, and in the case $N = 11$ the space $X_0(11)$ is particularly simple and has genus 1.

At this point we have just referred to two wildly different objects as $X_0(11)$. We will start from the Riemann surface and get the elliptic curve.

It is a fact from differential topology that the dimension of holomorphic differentials on a Riemann surface equals the genus. And, the holomorphic differentials on $X_0(N)$ are essentially the modular forms of weight 2; roughly speaking, there is a canonical isomorphism

$$f(z)dz \longrightarrow f(z).$$

The transformation law for modular forms is equivalent to the fact that these differentials are defined on the quotient $X_0(N)$.

Accordingly, $S_2(\Gamma_0(11))$ is one-dimensional. It's spanned by newforms (and the linear algebra of one-dimensional vector spaces is pretty trivial...) and we can write the unique normalized newform down:

$$f(z) = \eta(z)^2\eta(11z)^2 = q - 2q^2 - q^3 + \dots.$$

So the elliptic curve above will correspond to this modular form.

To see how, we choose $\tau_0 \in \mathbb{H}$ and define a function $F(z)$ by

$$F(z) = \int_{\tau_0}^{z} f(\zeta)d\zeta$$

as well as a related homomorphism $\Phi_f : \Gamma_0(N) \to \mathbb{C}$:

$$\Phi_f(\gamma) := \int_{\tau_0}^{\gamma(\tau_0)} f(\zeta)d\zeta$$

Here Φ_f turns out to be independent of τ_0. Moreover, the elliptic and parabolic elements of $\Gamma_0(N)$ (i.e., $|Tr\gamma| \leq 2$) will map to 0.

Returning to the case $N = 11$, one can calculate that

$$\Gamma_0(11) =< T := \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, V_4 := \begin{pmatrix} 8 & 1 \\ -33 & -4 \end{pmatrix}, V_6 := \begin{pmatrix} 9 & 1 \\ -55 & -6 \end{pmatrix} > .$$

We can therefore calculate the image of all of $\Gamma_0(11)$ under Φ. T is parabolic and hence maps to 0; it turns out that V_4 and V_6 map to complex numbers which are linearly independent over $\mathbb{R}$. They therefore give a lattice Λ in $\mathbb{C}$. The function F gives a map $\mathbb{H} \to \mathbb{C}$, and we compose with the quotient map to obtain a map $\mathbb{H} \to \mathbb{C}/\Lambda$. However, by our construction of Φ, F maps $\Gamma_0(11)$ into Λ and so we get a holomorphic map

$$\Gamma_0(11)\backslash\mathbb{H} \to \mathbb{C}/\Lambda$$

and **this gives us the elliptic curve we want.**

Now a priori we shouldn't necessarily be too surprised: we cooked up a lattice and declared that it was an elliptic curve. It turns out that (as Knapp says) two miracles occur: One, everything in sight will be defined over the rationals. It is clear (once one knows the relevant theory) that one can get an elliptic curve as above, but quite surprising that it should be defined over the rationals.

And the second miracle is, the L-functions will match. Why should this be....?

3. Sketch of Eichler-Shimura Theory

Theorem 3.1 (Eichler-Shimura)**.** *Assume that $f(z) = \sum_{n\geq 1} c_n q^n$ is a newform in $S_2(\Gamma_0(N))$ with coefficients in $\mathbb{Z}$. Then there exists an elliptic curve E associated to f satisfying the following:*

(1) E is defined over $\mathbb{Q}$, and E is a quotient of the Jacobian variety J by an abelian subvariety A which is also defined over $\mathbb{Q}$,

(2) The members $t(n)$ of $\mathrm{End}(J)$ leave A stable and act on E as multiplication by c_n,

(3) If one writes

$$\Lambda_f := \left\{ \Phi_f(\gamma) := \int_{\tau_0}^{\gamma(\tau_0)} f(\zeta) d\zeta \middle| \gamma \in \Gamma_0(N) \right\}$$

then Λ_f is a lattice in $\mathbb{C}$, with E isomorphic to $\mathbb{C}/\Lambda_f$ over $\mathbb{C}$,

(4) $L(E,s) = L(f,s)$.

If the coefficients of f are (algebraic integers) in some larger number field, then a similar statement is true. Here E will be an abelian variety of dimension equal to the degree of the number field over $\mathbb{Q}$. Importantly, we will still obtain a match of L-functions.

We will now attempt to describe what is meant, and talk about some of the ingredients that go into the proof.

3.1. **Definition of the L-functions.** We will review how the L-functions $L(E,s)$ and $L(f,s)$ are defined. First, the modular form. That's really, really easy. The L-function of a cusp form $f(z) = \sum_{n\geq 1} c_n q^n$ is just $L(f,s) = \sum_{n\geq 1} c_n n^{-s}$. We say that $L(f,s)$ is the **Mellin transform** of $f(z)$. "Mellin transform" means that everywhere you see q^n you write n^{-s} instead.

Really, the Mellin transform of a function f is

$$\int_0^\infty (f(y) - f(\infty)) y^{s-1} dy,$$

something that looks like the gamma function. Check the formula, you get what I claimed!

The L-function of an elliptic curve is follows. For a prime p of good reduction (all but finitely many primes are of good reduction) we define

$$a_p := p + 1 - \#E(\mathbb{F}_p).$$

For a prime p of bad reduction, we define a_p to be 1, -1, or 0 depending on the type of bad reduction. Then, we have

$$L(E,s) := \prod_p \frac{1}{1 - a_p p^{-s} + p^{1-2s}} \prod_p \frac{1}{1 - a_p p^{-s}}$$

where the first product is over the good primes and the second is over the bad.

There is a lot I could say about these kinds of L-functions and I won't say it here.

3.2. **A $\mathbb{Q}$-structure on $X_0(N)$.** We will use the theory of the j-function to show that $X_0(N)$ is "defined over" $\mathbb{Q}$ (while being somewhat vague about exactly what that means...).

Definition 3.2. *The modular function $j(z)$ is*

$$j(z) := \frac{E_4(z)^3}{\Delta(z)} = \frac{1}{q} + 744 + 196884q + \dots.$$

Here (and always) $q := e^{2\pi i z}$.

The j-function is holomorphic on $\mathbb{H}$, with a simple pole at infinity. We have the modular (of weight 0) relation $j(\gamma z) = j(z)$ for $\gamma \in SL_2(\mathbb{Z})$, so that j is a well-defined meromorphic function on $X_0(1)$.

Theorem 3.3. *Every meromorphic function on $X_0(1)$ (i.e, every modular function for $SL_2(\mathbb{Z})$ is a rational function in $j(z)$.*

Proof. (Sketch) We use the theory of Riemann surfaces here. First of all, we show that the j-function is 1-1 and onto (as a map from $Y_0(1)$ to $\mathbb{C}$. If a complex number z_0 is fixed, then the function $j(z)-z_0$ has exactly one pole in $X_0(1)$, hence it has exactly one zero.

We then show that the modular functions having poles only at infinity are polynomials in $j(z)$. To do this, we use the fact that the only holomorphic functions on $X_0(1)$ are constants. If $f(z) = a_{-1}q^{-1} + a_0 + a_1q^1 + \dots$ we see that it is a multiple of the j-function plus a constant. If $f(z)$ has lower-order terms in the Laurent series, we can use induction.

Finally, if $f(z)$ has a pole at some point p, then the function $(j - j(p))^{-1}$ has a simple pole at p, and no poles anywhere else. We subtract from f an appropriate combination of powers of this function, and in this way can get rid of any poles not at infinity. □

A nice way of rephrasing the above result is to say that

$$K(X_0(1)) = \mathbb{C}(j),$$

where $K(X_0(1))$ denotes the **function field** of $X_0(1)$ (i.e. the field of meromorphic functions).

Definition 3.4. *We define a function $j_N(z)$ (for a positive integer N) by*

$$j_N(z) := j(Nz).$$

We also define a **modular polynomial**

$$\Phi_N(X) := \prod (X - j \circ \alpha_i),$$

where the α_i are coset representatives for the Hecke operators, defined by

$$SL_2(\mathbb{Z}) \begin{pmatrix} N & 0 \\ 0 & 1 \end{pmatrix} \mathrm{SL}_2(\mathbb{Z}) = \cup SL_2(\mathbb{Z})\alpha_i.$$

By modularity of j, the definition of $\Phi_N(X)$ does not depend on the choice of coset representatives.

We note (**important!**) that j_N is a root of $\Phi(N)$, as j_N is just $j \circ \begin{pmatrix} N & 0 \\ 0 & 1 \end{pmatrix}$.

Theorem 3.5. *(1) The polynomial $\Phi_N(X)$ has coefficients in $\mathbb{Z}[j]$ and is irreducible over $\mathbb{C}[j]$. (2)*

$$K(X_0(N)) = \mathbb{C}(j, j_N).$$

Proof. (Brief sketch) The proof involves some Galois theory – here $K(X_0(N))$ is a field extension of $K(X_0(1))$, and we get do some Galois theory over $\mathbb{Q}$ too. $SL_2(Z)$ acts on the coefficients of $X_0(N)$, and one shows it just permutes the roots. There are a number of other details involved, which I skip here. □

What does this theorem say? The function field of $X_0(N)$ has transcendence degree 1, and is generated by two elements which satisfy an algebraic relation with coefficients in $\mathbb{Z}$.

By analogy, consider some elliptic curve $E : y^2 = x^3 + ax + b$. Then the function field of E is (for whatever field k) simply $k(x, y)$, where x and y satisfy the relation given by the equation for E. If a and b are rational, then E is "defined over $\mathbb{Q}$", and this is detected by the function field.

The idea is that algebraic geometry gives a dictionary between curves and their function fields. We will skip any discussion of this that is at all precise. (But, see Knapp.) The point is that through this dictionary and the theorem given above, we have given a canonical $\mathbb{Q}$-structure to $X_0(N)$.

3.3. **Hecke Operators.** The *Hecke operators* form an algebra of operators on the spaces of modular forms $M_2(\Gamma_0(N))$, and of cusp forms $S_2(\Gamma_0(N))$. They have a lot of different descriptions – for example in terms of lattices or double cosets. They can also be given a simpler (but less intrinsic) description in terms of the Fourier coefficients of the modular forms involved. (See, e.g, p. 21 of Ken's book.)

We will not go into them in much detail here. However we will discuss a couple of key features. For one, it turns out that the Hecke operators all commute. It then follows from linear algebra that $S_2(\Gamma_0(N))$ has a basis of simultaneous eigenforms for all of the Hecke operators. One can break this space up further into *newforms* and *oldforms* , the oldforms being those modular forms that "come from" forms of lower level.

What is especially important for the Eichler-Shimura theory is, as Knapp says: "Hecke operators act on many things, and they do so consistently. Understanding these consistent actions is a key to unlocking the power of the Hecke operators."

We will discuss Hecke operators from the standpoint of differential topology here. We will give only give an overview here (because... ahem... the author of this paper had only a few weeks and did not make it through all of the details...)

Hecke operators on homology. The Riemann surface (quotient space, modular curve, etc.) $X_0(N)$ has a certain *genus* g. The genus of a Riemann surface is a subtle and powerful invariant that tells you a lot of useful things. For one, it is the dimension of the space of holomorphic differential forms. As noted before, holomorphic differentials on $X_0(N)$ are essentially the same thing as modular forms for $\Gamma_0(N)$. (The spaces are canonically isomorphic, with the isomorphism being "write down a dz after the modular form".)

The genus also tells you how many holes $X_0(N)$ has as a complex donut. From algebraic topology, we obtain the following

Proposition 3.6.

$$H_1(X_0(N), \mathbb{Z}) \simeq \mathbb{Z}^{2g}.$$

However, we also obtain the following interesting canonical isomorphism:

Theorem 3.7.

$$H_1(X_0(N), \mathbb{Z}) \simeq \Gamma_0(N)^{ab}/\Gamma_{ep}^{ab}.$$

Here G^{ab} is the abelianization of G (G modulo its commutator subgroup), and Γ_{ep} is the subgroup generated by all the elliptic and parabolic elements.

Proof. (Sketch) The basic idea is that morally (but not literally) speaking, $X_0(N)$ is the quotient of $\mathbb{H}$ by an action of $\Gamma_0(N)$. One uses the Galois correspondence between covering spaces and subgroups of the fundamental group to conclude that $\pi_1(X_0(N)) \simeq \Gamma_0(N)$. To pass from fundamental group to homology, one then just abelianizes.

The above is not literally true because the action of $\Gamma_0(N)$ has fixed points, corresponding to the elliptic and parabolic elements. So, one first removes them and obtains essentially the above isomorphism. Then, we add the fixed points back. In doing this, we consider small loops around them, and when the fixed points are added these loops become trivial in $\pi_1(X_0(N))$. The effect of this is to quotient $\pi_1(X_0(N))$ by relations corresponding to each of the fixed points, and therefore each of the elliptic and parabolic elements. One checks carefully that this corresponds to quotienting by Γ_{ep}, and passing to the abelianization the result follows. □

There is also a correspondence between elements of $\Gamma_0(N)$ and cycles on $X_0(N)$. One chooses a base point z_0, and then an element γ corresponds to a path from z_0 to γz_0, which in turn corresponds

to a loop in $X_0(N)$. (Once again, there are arbitrary choices involved, and one is obliged to check that they don't matter.)

So what's the point? Given all these correspondences, one sees that the Hecke algebra **acts on everything in sight.** And it does so compatibly, as shown by the following proposition.

Proposition 3.8. *The Hecke operators $T(n)$ act as $\mathbb{Z}$-linear operators on $H_1(X_0(N), \mathbb{Z})$. Moreover, for a 1-cycle c on $X_0(N)$ and a holomorphic differential ω on $X_0(N)$ (which corresponds to some cusp form) we have*

$$\int_{T(n)c} \omega = \int_c T(n)\omega.$$

This is not immediate, but it's not exactly deep either. The hard part is getting the definitions right. (We are doing topology after all...)

3.4. **Abelian varieties.** An **abelian variety** A is a nonsingular projective variety over $\mathbb{C}$ with a distinguished point O, and an abelian group structure such that O is the identity and addition and subtraction are morphisms.

One can define quotient varieties. Suppose A is an abelian variety and C is an abelian subvariety. Then there exists an abelian variety, unique up to canonical isomorphism, which we will call A/C, and a surjective homomorphism $A \to A/C$ whose kernel is exactly C. Moreover, every morphism $A \to B$ whose kernel contains C factors through A/C.

3.5. **The Jacobian variety.** In general, assume that X is a Riemann surface of genus g. Then we have $H_1(X, \mathbb{Z}) = \mathbb{Z}^{2g}$. Let $c_1, \dots, c_{2g}$ be a $\mathbb{Z}$-basis for this homology group, and let $\omega_1, \dots, \omega_g$ be a basis for the space of holomorphic differentials on X. Then, the $2g$ vectors

$$\Big(\int_{c_k} \omega_1 \dots \int_{c_k} \omega_g\Big)^T$$

in $\mathbb{C}^g$ are linearly independent over R. Therefore, they form a $\mathbb{Z}$-basis for a lattice $\Lambda(X) \in \mathbb{C}^g$.

Definition 3.9. *The Jacobian variety of X is the g-dimensional complex torus $J(X) := \mathbb{C}^g/\Lambda(X)$. Moreover, if $x_0 \in X$ is fixed we have an injective holomorphic map $\Phi : X \to J(X)$ given by*

$$\Phi(x) = \left\{ \int_{x_0}^{x} \omega_j \right\}_{j=1}^{g}.$$

Moreover, $J(X)$ is an abelian variety, with the group law the one induced by addition in $\mathbb{C}^g$.

If the genus is 1, this map is surjective (this is the elliptic curve case). If the genus is larger, it isn't.

We will go ahead and specialize to the case of interest, $X = X_0(N)$, at this point. Write $J = J(X_0(N))$ for its Jacobian variety. By the definition, the endomorphism ring $End(J)$ is a subring of the ring $M(g, \mathbb{C})$ of g-by-g complex matrices. We wish to realize the Hecke operators as members of $End(J)$ and hence as g-by-g matrices.

We start with the action of the Hecke alegbra on the divisor group of $X_0(N)$. The n-th Hecke operator is defined here by

$$T(n)([z]) = \sum_i [\alpha_i z],$$

where we have written $M(n, N)$ as the set of 2-by-2 integer matrices of determinant n, whose lower left entry is prime to N, and whose lower left entry is divisible by N. We have then chosen a coset decomposition

$$M(n, N) = \cup \Gamma_0(N)\alpha_i.$$

As always, one is obliged to check that this is well-defined.

We recall that we have a map $X_0(N) \to J$, and this also induces a map $\mathrm{Div}(X_0(N)) \to J$. We therefore obtain a second map $X_0(N) \to \mathrm{Div}(X_0(N)) \to J$, where the first map is the Hecke operator $T(n)$.

At this point, we have two maps $X_0(N) \to J$, and we can invoke the universal mapping property for abelian varieties. We therefore obtain a map $J \to J$, which we will call $t(n)$, and an appropriate commutative diagram.

There are further details which I do not pursue here, but this is the construction which realizes the Hecke algebra as endomorphisms of the Jacobian. We have the following proposition:

Proposition 3.10. *Choose a basis $\{f_1, \ldots, f_g\}$ of $S_2(\Gamma_0(N))$ over $\mathbb{C}$ so that the Hecke operators are all given as g-by-g integer matrices. (See pp. 327-329 of Knapp for a proof that it is possible to do this.) Also consider the differentials $dt(n)$ of the endomorphisms $t(n)$ constructed above, also regarded as g-by-g matrices as above.*

Then we have as matrices $T(n) = dt(n)$.

3.6. **Overview of the proof.** We start by looking at the Hecke algebra T. We look at the endomorphisms $t(n) \in \mathrm{End}(J)$ and tensor with $\mathbb{Q}$; we may identify each $t(n)$ with its differential $dt(n)$, and so T will be a subalgebra of the algebra $M_g(\mathbb{Q})$ of g-by-g matrices.

We now invoke the Wedderburn theorem, and we obtain

$$T = k_1 \oplus k_2 \oplus \cdots \oplus k_r \oplus R,$$

where R is the nilradical, and each k_i is isomorphic to a finite extension of $\mathbb{Q}$.

We now define a homomorphism $\rho : T \to \mathbb{Q}$ by using the previous proposition. In particular, each element of T can be identified with its eigenvalue for the modular form f; i.e., its n-th Fourier coefficient c_n. Therefore we define

$$\rho(t(n)) := c_n$$

and by changing the order of the factors if needed we see that $\rho(R) = 0$ and $\rho(k_1) = \mathbb{Q}$. Therefore $k_1 \simeq \mathbb{Q}$. We define an ideal U of T by

$$U = k_2 \oplus \cdots \oplus k_r \oplus R.$$

Now we define $A \subset J$ to be the abelian subvariety consisting of $\alpha(J)$, as α runs over all endomorphisms in U. Each α is a homomorphism of abelian varieties, and as one can (nontrivially) check, it is defined over $\mathbb{Q}$. We also check that the sum of two abelian subvarieties of J will again be an abelian subvariety, and it will be defined over $\mathbb{Q}$ if the summands are. So, as we range over α we keep making A larger.

We can therefore take the quotient $E := J/A$. We need to check that A has codimension exactly 1. We will vaguely describe the argument. In the first place, we need to show that $A \neq J$. To do this, we use the algebraic characterization given earlier, and exhibit a nonzero element of $\mathrm{End}(J)$ that annihilates A. To show that $\dim E \leq 1$, we suppose to the contrary that there are two linearly independent holomorphic differentials ω and ω' on E. We chase down our definitions and use these to define two linearly independent newforms f and f' which have the same set of Hecke eigenvalues. But this is impossible due to the "multiplicity one" phenomenon.

3.7. **Match of L-functions.** There is also the match of L-functions to justify. This is the most difficult part of the argument, and at this point I will become even more vague. The general argument involves a substantial amount of algebraic geometry in characteristic p, which is well beyond the expertise of the author (as well as the scope of Knapp's book).

We will (following Knapp) at least say something about the case where $X_0(N)$ has genus 1. In this case, part (2) of the main theorem says essentially that the reduction of the Hecke operator modulo p is essentially $[c_n]$, the nth Fourier coefficient.

On the other hand, we have a useful formal identity within $\mathrm{End}(E_p)$. (Here E_p is the reduction of E modulo p.) We write ϕ for the *Frobenius morphism* sending all coefficients of E to their pth powers. It has a *dual isogeny* $\hat{\phi}$, which satisfies $\hat{\phi} \circ \phi = [p]$, the multiplication by p map.

We start with the equation

$$[\#E(\mathbb{F}_p] = [deg([1] - \phi)].$$

Essentially, this says that the points in $E(\overline{\mathbb{F}_p})$ which are in fact $E(\mathbb{F}_p)$ are exactly the fixed points of the Frobenius map, so that they form the kernel of the isogeny $[1] - \phi$.

Now we have $[deg([1] - \phi)] = ([1] \hat{-} \phi) \circ ([1] - \phi)$ so that we can just FOIL everything. We obtain

$$[\#E(\mathbb{F}_p)] = [1] - (\phi + \hat{\phi}) + \hat{\phi} \circ \phi = [1] - (\phi + \hat{\phi}) + [p]$$

and upon rearranging,

$$\phi + \hat{\phi} = [p + 1 - \#E(\mathbb{F}_p)] = [a_p].$$

The idea, then, is to prove that

$$\phi + \hat{\phi} = [c_p]$$

and then we get $a_p = c_p$, which is what we want.

So we have at least managed to state the result we want to prove, for a special case (genus 1) of a special case (coefficients in $\mathbb{Z}$) of the Eichler-Shimura theory. This is far as Knapp goes, and regrettably this will be all that I will say as well.

4. Existence of Rank 0 Quotients

In his proof of Corollary 4.4, Mazur needs the following fact:

Theorem 4.1. *The Jacobian $J_0(N)$ associated to the modular curve $X_0(N)$ has a quotient abelian variety A whose rank is zero.*

To prove this, we will establish the following fact:

Theorem 4.2. *(Duke) For appropriate N, there is a Hecke eigenform $f(z) \in S_2(\Gamma_0(N))$ such that $L(f, 1) \neq 0$.*

Why does this help us? The Eichler-Shimura theory discussed above (or more properly, a generalization of it) tell us that the Jacobian $J_0(N)$ breaks up into abelian quotients (e.g., elliptic curves) A associated to the cusp forms for $\Gamma_0(N)$. In particular, if A is the abelian variety associated to $f \in \Gamma_0(N)$, then we have $L(A, s) = L(f, s)$.

So, we have a quotient variety A whose L-function does not vanish at the critical value. The Birch and Swinnerton-Dyer conjecture (yes, there's a more general version for abelian varieties...) then implies that the rank of A is 0.

Happily, BSD is now known in the relevant case! (It was not known at the time Mazur wrote his paper.) By work of Kolyvagin (is this the work we have been studying...? it seems like it isn't), it is known that if $L(A, s) = 0$ then A has rank 0, in the cases where A is an elliptic curve or else has "real multiplication".

The rest of this talk will given an overview of the first result.

Remark: The theorem attributed to Duke above was actually known before Duke; there was a prior proof due to Darmon and Merel using the theory of modular symbols. However, there is a nice quantitative version of Duke's theorem, Duke's theorem applies more generally, and the proof is fairly easy (if one is willing to accept some facts about Poincare series.) So we will talk about Duke's approach here.

5. Duke's theorem on nonvanishing of modular L-functions

Let F be a basis for the weight 2 cusp forms in $S_2(\Gamma_0(N))$. We will assume throughout that F has been chosen to be orthonormal with respect to the Petersson inner product. (Duke normalizes instead so that the first Fourier coefficient is 1; he starts with a different version of the Petersson trace formula and the proof is the same.)

Duke's result is the following:

Theorem 5.1. *(Duke)*

$$\sum_{f\in F} a_1(f)L(f,1) = 4\pi + O(N^{-1/2}\log N).$$

In other words, for sufficiently large N, a weighted average of the central critical values is nonzero. Thus (duh...) at least ONE of them is not zero.

The same is true when the modular forms f are twisted by a Dirichlet character χ, so that the sum is over the L-values $L(f\otimes\chi,1)$. We won't concern ourselves with this case.

The error term was improved and made explicit by Ellenberg:

Theorem 5.2. *(Ellenberg)*

$$\sum_{f\in F} a_1(f)L(f,1) = 4\pi\exp(-2\pi/\sigma N\log N) + O(N^{-1+\epsilon}).$$

This is true for a certain range of real values σ. Note that the exponential goes to 1 with N. This is the 'simple' version of the theorem; the main theorem of the paper gives an explicit upper bound for the error. In particular, the total error is bounded by a sum of five error terms, each of which is complicated. One of them involves the number $(400/399)^3$.

Although we will not go through the entire proof, we will outline several of the key ingredients.

5.1. The Petersson trace formula.

The Petersson trace formula says,

$$\frac{1}{4\pi\sqrt{mn}}\sum_{f\in F} a_m(f)a_n(f) = \delta_{mn} - 2\pi \sum_{c>0;c\equiv 0 \mod N} c^{-1}S(m,n;c)J_1(4\pi\sqrt{mn}/c).$$

Here $S(m,n;c)$ is a *Kloosterman sum* , and J_1 is the J-Bessel function. If you have to ask, then you probably don't want to know.

This is the formula for weight 2, where the Fourier coefficients of our modular forms will be real. When this is not assumed to be the case, something only slightly more complicated is true.

What is this formula says is that the Fourier coefficients of modular forms are *almost orthogonal.* One uses standard bounds on the stuff on the right to show that it is $\delta_{mn} + o_N(1)$ (where $o_N(1)$ is something explicit).

Before showing how to apply this formula, we will sketch the proof. (For this proof, we were referring to Iwaniec and Kowalski's book, and so we are stating our results for k not necessarily 2.)

The m-th **Poincare series** is

$$P_m(z) := \sum_{\gamma\in\Gamma_\infty\backslash\Gamma_0(N)} (cz+d)^{-k}\exp\left(2\pi i m\frac{az+b}{cz+d}\right).$$

If $m = 0$ you just get the Eisenstein series of weight k. But if $m > 0$ you still get modular forms! (and cusp forms even). In fact we have the following Fourier expansion:

Proposition 5.3. *If $m > 0$ then we have*

$$P_m(z) = \sum_{n\geq 1} p(m,n)q^n$$

where

$$p(m,n) = (m/n)^{(k-1)/2}\left(\delta_{mn} + 2\pi i^{-k} \sum_{c>0; c\equiv 0 \mod N} c^{-1} S(m,n;c) J_{k-1}(4\pi\sqrt{mn}/c)\right).$$

Look familiar? Substitute in $k = 2$ and you get exactly the thing above.

To prove the proposition, you just kind of write out everything and see what you get. There are a lot of details; the Bessel functions arise from Poisson summation, etc.

The **Petersson inner product** is defined by

$$< f, g >:= \int f(z)\overline{g(z)} y^k d\mu(z)$$

where the integration is done over a fundamental domain for $\Gamma_0(N)$ and $\mu(z) := dx\ dy/y^2$ represents the invariant measure. One checks that the integral is well-defined (i.e., is $\Gamma_0(N)$-invariant).

One then has the following

Proposition 5.4. *If $f \in M_k(\Gamma_0(N)) = \sum_{n\geq 0} a_f(n) q^n$, then for any $m \geq 1$*

$$< f, P_m >= \frac{\Gamma(k-1)}{(4\pi m)^{k-1}} a_f(m).$$

This follows by a surprisingly simple calculation. One then obtains the following

Corollary 5.5. *The Poincare series with $m \neq 0$ span $S_k(\Gamma_0(N))$.*

Proof. If not, there is a modular form which is orthogonal to all of the Poincare series... now use the above formula, all the Fourier coefficients are zero. □

So what can we use the above formula to prove? Given a basis f of cusp forms of level N, which we will assume has been chosen to be orthonormal with respect to the Petersson product, we write

$$P_m = \sum_f < f, P_m > f$$

and use the above proposition to obtain

$$P_m = \frac{\Gamma(k-1)}{(4\pi m)^{k-1}} \sum_f a_f(m) \sum_n a_f(n) q^n.$$

Comparing this with the explicit formula for $p(m,n)$ we obtain the Petersson trace formula.

5.2. **The proof of Duke's theorem.** We will simplify the Petersson trace formula, and write

$$\frac{1}{4\pi\sqrt{mn}} \sum_{f\in F} a_f(m) a_f(m) = \delta_{mn} + o(1).$$

We won't worry too much about what $o(1)$ means.

The starting point is the following equation for the central critical value:

$$L(f,1) = \sum_n a_f(n) \frac{1}{n} e^{-2\pi n/x} + \epsilon \sum_n a_f(n) \frac{1}{n} e^{-2\pi n x/N}.$$

This formula appeared in HW 11 in elliptic curves last year. Here ϵ is the root number (i.e. it is *not* an arbitrarily small number) and x is an arbitrary real number.

We choose $x = N \log N$ and obtain the formula

$$L(f,1) = \sum_n a_f(n) \frac{1}{n} e^{-2\pi n/(N\log N)} \pm \sum_n a_f(n) \frac{1}{n} e^{-2\pi n \log N}.$$

Now we have the Ramanujan bound

$$|a_f(n)| \leq d(n)n^{1/2},$$

which means the second part of the sum converges REALLY fast. (Duke writes $O(N^{-6})$. As I said...) We could get faster convergence with a choice like $x = \sqrt{N}$, but since we don't know what the root number is, we want to minimize that part of the sum.

So, we obtain the formula

$$\sum_{f \in F} a_f(1)L(f,1) = \sum_f a_f(1) \sum_n a_f(n) \frac{1}{n} e^{-2\pi n/(N \log N)} + o(\dots).$$

Now the Petersson trace formula with $m = 1$ tells us that

$$\frac{1}{4\pi\sqrt{n}} \sum_f a_f(n) = \delta_{n,1} + o(1).$$

So if $n = 1$ then $\sum_f a_f(1)a_f(n)$ is very nearly 4π. For larger n we get something really small. We have to sum over an infinite range of n, but we are summing against exponential decay, so it is not surprising that we again get $o(1)$.

We finally obtain

$$\sum_{f \in F} a_f(1)L(f,1) = 4\pi e^{-2\pi/N \log N} + o(1).$$

As N gets large the exponential is basically 1, and the result follows.

References

[1] W. Duke, *The critical order of vanishing of automorphic L-functions with large level*, Invent. math., 119 (1995), 165-174.

[2] J. Ellenberg, *Galois representations attached to $\mathbb{Q}$-curves...* (see his website)

[3] J. Ellenberg, *On the error term in Duke's estimate for the average special value of L-functions* (see his website)

[4] A. Knapp, *Elliptic curves*, Princeton University Press, Princeton, 1992. **It's really good, read it!**

[5] Ken, *Ken's book*.

Department of Mathematics, University of Wisconsin, Madison, Wisconsin 53706

E-mail address: thorne@math.wisc.edu