AN APPLICATION OF THE EICHLER-SHIMURA CONGRUENCE RELATION

ANDREAS HATZIILIOU

1. Prelude

In this short expository paper, I aim to give a brief overview, as well as some insight, regarding an area of mathematics which I find particularly pleasant. This being the theory of Hecke operators and their role in various areas of number theory and algebra. I will begin this section by laying out a few necessary and relevant definitions in order to motivate our discussion.

Definition. *Let K be an algebraically closed field, let P^n be a projective space of dimension n over K. Let $f \in K[x_1, \ldots, x_n]$ a homogeneous polynomial of degree d. For each set of homogenous polynomials S we define the zero-locus*

$$Z(S) = \{x \in P^n : f(x) = 0 \forall f \in S\}$$

We define a projective variety V to be such a subset of P^n such that $V = Z(S)$ for some S and it irreducible (i.e. isn't the union of two proper subsets).

Definition. *An abelian variety A is a projective variety with morphisms*

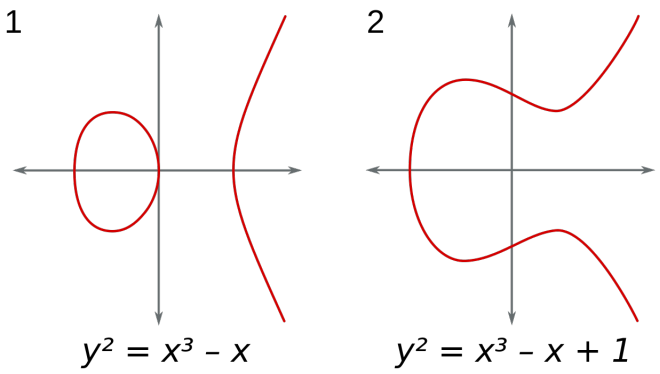
$$a : A \times A \to A$$
$$b : A \to A$$

along with an "identity" element e which induces a group structure on the variety.

There are different, equally interesting, ways of thinking about abelian varieties:

1 If the underlying field $K = \mathbb{C}$ and $\dim(A) = \mathrm{g}$, then we can think of them complex torus (with the structure of a complex manifold) of dimension g that is also a projective variety over $\mathbb{C}$.
2 In a categorical view-point, an abelian variety A is a group in the category of projective varieties.

A simple example of the latter are *elliptic curves*, which are abelian varieties of dimension 1. This tells us both of the curves below are of genus 1!



Definition. *Let R be a commutative ring. The special linear group $SL_n(R) \leqslant GL_n(R)$ is such that $SL_n(R) := \{g \in GL_n(R) : det(g) = 1\}$*

Definition. *For some $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{R})$, some $z \in \mathbb{C}$ we define the "fractional linear transformation"*

$$gz \mapsto \frac{az+b}{cz+d}$$

and moreover g send the point at infinity to $\lim_{z\to\infty} gz = \frac{a}{c}$. This type of transformation defines a group action on $\tilde{\mathbb{C}}$.

It is readily shown that the two only matrices in $SL_2(\mathbb{R})$ which act trivially on $\tilde{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ are $\pm I$. Thus, by quotient out by $\{\pm I\}$ we obtain a group which acts faithfully on $\mathbb{C}$. This group acts by transformations on the upper half plane $\mathbb{H}$.

Definition. *This construction leads us to define to be the "Modular Group", a group of central importance in number theory. The full modular group is defined as $\Gamma = SL_2(\mathbb{Z})$, and is the group of all rational fractional linear transformations. Moreover we denote $\bar{\Gamma} = \Gamma/\{\pm I\}$, which acts faithfully on $\mathbb{H}$.*

We now introduce a few important subgroups of the modular group which will be useful for the coming sections

Definition. *Let $N \in \mathbb{N}$. We introduce a normal subgroup of $SL_2(\mathbb{Z})$ called the "principal congruence subgroup of level N" to be*

$$\Gamma(N) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}) \middle| \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \mod N \right\}$$

We can see this subgroup to be normal as it is the kernel of the map $\phi : \Gamma \to SL_2(\mathbb{Z})$ sending $g \mapsto g \mod N$. We call a subgroup of Γ a congruence subgroup of level N if it contains $\Gamma(N)$. Remark that any congruence subgroup of level N will also contain level M when M is a multiple of N.

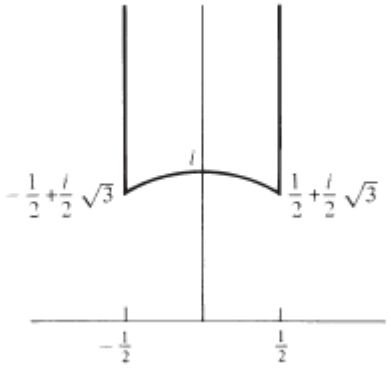
In the context of this paper, we will often work with the following two congruence subgroups of Γ:

$$\Gamma_0(N) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma \middle| c \equiv 0 \mod N \right\}$$

$$\Gamma_1(N) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N) \middle| a \equiv 0 \mod N \right\}$$

We now consider the group action of $\Gamma \circlearrowright \mathbb{H}$ in order to define its fundamental domain.

The fundamental domain of Γ is $F := \{z \in \mathbb{H} : |Re(z)| \leqslant \frac{1}{2} and |z| \geqslant 1\}$



Fundamental domain of $SL_2(\mathbb{Z})$

We now briefly discuss an alternative and more convenient way of working with modular forms. Consider the map from $\bar{\mathbb{H}} = \mathbb{H} \cup \{\infty\}$ to the punctured unit disc given by

$$z \mapsto q = e^{2\pi i z}$$

where we map the point at infinity in $\mathbb{H}$ to the origin. We then get Laurent expansions for our complex functions which resemble

$$f(z) = \sum_{n \in \mathbb{Z}} a_n q^n$$

This allows us to talk about meromorphicity and holomorphicity at ∞ by stating that:

- f is meromorphic at ∞ if the q-expansion has only finitely many nonzero a_n with $n < 0$
- f is holomorphic at ∞ if the q-expansion has no nonzero a_n with $n < 0$
- f is said to vanish at ∞ if it is holomorphic on $\mathbb{H}$ and $a_0 = 0$.

Definition. *Let $k \in Z$, let $f(z)$ be meromorphic on $\mathbb{H}$ with q-expansion $\sum_{n \in \mathbb{Z}} a_n q^n$. Suppose $f(z)$ satisfies*

$$f(z) = (cz + d)^k f(z) \quad \forall = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z})$$

(1) If $f(z)$ is meromorphic at infinity then we call f a "modular function" of weight k for Γ.
(2) If $f(z)$ is holomorphic on $\mathbb{H}$, then we call f a "modular form" of weight k for Γ.
(3) Finally if f additionally has $a_0 = 0$, then we call f a "cusp form" of weight k for Γ

We denote the space of all modular forms of weight k for Γ by $M_k(\Gamma)$. The space of cusp forms is denoted $S_k(\Gamma)$.

We now want to generalize the definition to congruence subgroups. Let $GL_2^+(\mathbb{Q})$ be the subgroup of $GL_2(\mathbb{Q})$ consisting of matrices of positive determinant. Let $f(z)$ be a function on

$\bar{\mathbb{H}}$, let $k \in \mathbb{Z}$ and $\gamma \in GL_2^+(\mathbb{Q})$. We denote

$$f(z)|[\gamma]_k := (\det \gamma)^{k/2}(cz+d)^{-k} f(\gamma z) \qquad \text{for} \qquad \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Given this notation, it is clear that any modular function of weight k for Γ satisfies $f|[\gamma]_k = f$ for any $\gamma \in \Gamma$. This now allows us to define modular functions, forms and cusp forms for the congruence subgroups $\Gamma' \supset \Gamma(N)$.

If $f(z)$ is a meromorphic function on $\mathbb{H}$, $\Gamma' \subset \Gamma$ a congruence subgroup of level N, $k \in \mathbb{Z}$. If

$$f|[\gamma]_k = f \quad \text{ for all } \quad \gamma \in \Gamma'$$

and for any $\gamma_0 \in \Gamma$ $f(z)$ has a q-expansion $\sum a_n q_N^n$ with only finitely many non-zero a_n with $n < 0$, we call f a modular function of weight k for Γ'. We define similarly as before, a modular form of weight k on Γ', to be an f which is holomorphic on $\mathbb{H}$ and has no non-zero a_n for $n < 0$. Again, cusp forms have the additional requirement of vanishing at ∞. We denote the space of all modular forms of weight k for Γ' by $M_k(\Gamma')$. The space of cusp forms is denoted $S_k(\Gamma')$.

Modular forms have been of particular interest to me in the last year as they seem to appear (and be applicable to) a wide range of mathematical topics. A particularly fun and well-known instance of this type of behaviour is in the case of the *Monster Group* being related to the j-invariant. We define the j-invariant (or j function) as

$$j(z) = \frac{1728 g_2(z)^3}{\Delta(z)}$$

where

$$g_2(\tau) = 60 \sum_{(m,n)\neq(0,0)} (m+n\tau)^{-4}$$

is the function which appears in the laurent expansion of the Wierstrass $\wp$ function and $\Delta(z)$ is the modular discriminant. Now, the story goes that as John McKay was studying the j-invariant, he wrote out the first few fourier coefficients

$$j(\tau) = \frac{1}{q} + 196884q + 21493760q^2 + 864299970q^3 + 20245856256q^4 + \cdots$$

and realized that there was a surprising connection between them, and linear combinations of the dimensions of irreducible representations of the monster group M. The monster group is the largest sporadic simple group (it has order 808,017,424,794,512,875,886,459,904,961,710,757,005 ,754,368,000,000,000). The first few relations are as follow:

Dim. of irreducible representations of M	Coeff. of $j(\tau)$	Relation
$r_1 = 1$	1	$1 = r_1$
$r_2 = 196883$	196884	$196884 = r_1 + r_2$
$r_3 = 21296876$	21493760	$21493760 = r_1 + r_2 + r_3$

2. Hecke operators

First, to give some historical context, it is interesting to note that Hecke operators were studied (almost) simultaneously by L.J. Mordell and E. Hecke. Mordell used them in his 1917 paper proving the weak multiplicativity of fourier coefficients of the τ function stating that $\tau(ab) = \tau(a)\tau(b)$ when a, b are coprime. The theory was later generalized and made more complete by Hecke in his 1937 paper "Über Modulfunktionen und die Dirichletschen Reihen mit Eulerscher Produktentwicklung. I.". The study of Hecke operators has led us to some extremely robust machinery and has played a pivotal role in many of the "big" theorems of number theory and algebra proven in the last few decades.

Now, in more modern times, we have a few different ways of "building up" to the theory of Hecke operators. One of these uses the idea of modular point, but this doesn't generalize well in situations where no equivalent to modular points exists. Hence, we motivate our following definitions by noting that it is applicable to a much more general set of spaces

Definition. *Let A and B be groups, we say that they are commensurable if their intersection $A \cap B$ has finite index in both A and B.*

Proposition 2.1. *Let $H \subset G$ be any subgroup, $a \in G$ any element such that H and $a^{-1}Ha$ are commensurable. Let $H' = H \cap a^{-1}Ha$. Let $[H : H'] = d$ and write $H = \bigcup_{j=1}^{d} H'h_j$. Then $HaH = \bigcup_{j=1}^{d} Hah_j$ is a disjoint union of d right cosets. Conversely, if $HaH = \bigcup_{j=1}^{d} Hah_j$ is a disjoint union of d right cosets, then $H = \bigcup_{j=1}^{d} H'h_j$*

We define Hecke operators on a class of congruence subgroups of Γ which includes $\Gamma_0(N)$, $\Gamma_1(N)$, $\Gamma(N)$. Let $S^+ = m\mathbb{Z}$ for some positive integer m, $S^\times$ be a subgroup of $(\mathbb{Z}/N\mathbb{Z})^\times$, or alternatively denote the subset of $\mathbb{Z}$ whose image mod N is in $S^\times$.

$$\Delta^n(N, S^\times, S^+) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL_2(\mathbb{Z}) : N|c, a \in S^\times, b \in S^+, \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = n \right\}$$

Here are some examples:

$$\Gamma_1(N) = \Delta^1(N, 1, \mathbb{Z}) \quad \Gamma_0(N) = \Delta^1(N, (\mathbb{Z}/N\mathbb{Z})^\times, \mathbb{Z}) \quad \Gamma(N) = \Delta^1(N, 1, N\mathbb{Z})$$

Definition. *Let Γ' be a congruence subgroup of Γ and $\alpha \in GL_2^+(\mathbb{Q})$. Let $\Gamma'' = \Gamma' \cup \alpha^{-1}\Gamma'\alpha$ and $d = [\Gamma' : \Gamma'']$, $\Gamma' = \bigcup_{j=1}^{d} \Gamma''\gamma_j'$. Let $f(z)$ be a function on $\mathbb{H}$ invariant under $[\gamma]_k$ for $\gamma \in \Gamma'$. Then*

$$f(z)|[\Gamma'\alpha\Gamma']_k := \sum_{j=1}^{d} f(z)|[\alpha\gamma_j']_k$$

Note that this definition does not depends on the choice of representative α nor on the choice of the representatives γ_j' of Γ' mod Γ''. If $f \in M_k(\Gamma')$ it follows that $f(z)|[\Gamma'\alpha\Gamma']_k \in M_k(\Gamma')$.

Definition. *Let $\Gamma' = \Delta^1(N, S^\times, S^+)$ and let n be a positive integer. Let $f \in M_k(\Gamma')$. Then, we define a Hecke operator T_n to act as*

$$T_n f := n^{k/2-1} \sum f|[[\Gamma'\alpha\Gamma']_k]$$

where the sum is taken over all double cosets of Γ' *in* $\Delta^n(N, S^\times, S^+)$. *By the previous proposition, we have* $T_n f \in M_k(\Gamma')$

3. The Eichler-Shimura congruence relation

We begin this section by stating a wonderful theorem of Shimura in [2]. This theorem is a generalization of the work in his 1958 paper "Correspondances modulaires et les fonctions ζ de courbes algébriques" which relates the L-function of an abelian variety A induced by some modular form f to the $L-$function of f. The reason why this is interesting is that it gives us certain insights regarding the relationship between the geometric objects (varieties) and number-theoretic objects (modular curves, modular forms). This is the focus of the Langlands programme and we will hopefully see many more unexpected connections between the the major fields of math in years to come!

Theorem 3.1. *Let* $f \in S_2(\Gamma_0(N))^{new}$ *a normalized new-form which is a common eigen-function of each operator* T_n *for all* n. *Let* $f|T_n = a_n f$ *and let* K_f *be a subfeild of* $\mathbb{C}$ *generated by adjoining* $\mathbb{Q}(a_n : n \in \mathbb{N})$. *Then, there exists an abelian variety* A *and an isomorphism* $\theta : K_f \to End_{\mathbb{Q}}(A)$ *such that*

1 $\dim(A) = [K_f : \mathbb{Q}]$ and
$$Jac(X_0(N)) \twoheadrightarrow A$$
is a surjective morphism over $\mathbb{Q}$.
2 The hecke operators T_n of $\mathrm{End}(Jac(X_0(N)))$ act on A as mutliplication by the coefficients a_n. Hence $\theta(a_n)$ is the restriction of T_n acting on the Jacobian.
3 The couple (A, θ) is uniquely determined by (1) and (2) up to isomorphism. Moreover, for every isomorphism σ of K into $\mathbb{C}$, there exists an element $f_\sigma \in S_2(\Gamma_0(N))$ such that forall n
$$f_\sigma|T_n = a_n^\sigma f_\sigma$$
Moreover, this lets us write the L-function of our variety in terms of the L-function of the newform f
$$L(A/\mathbb{Q}, s) = L(f, s)$$

Now, I will go over a *very* bare-bones sketch of a proof in order to cover a few of the main ideas as well as introduce the Eichler-Shimura congruence relation. This sketch was inspired by the paper here[4].

1. We begin by considering modular curves $Y(\Gamma')$. These are inherently geometric objects which are realized in the following way:

- We start with some congruence subgroup $\Gamma' \subset SL_2(\mathbb{Z})$.
- Quotient the upper half plane $\mathbb{H}$ by the action of Γ'.
- When topologized, the result $\Gamma'\backslash\mathbb{H}$ is endowed with the structure of a Riemann Surface.
- We can compactify it by adding all the cusps of Γ' to obtain $X(\Gamma')$ and so this is a compact Riemann Surface (hence a projective curve).

2. There is a relationship between the latter and modular forms. Using the usual notation of differential geometry and letting $A_k(\Gamma)$ denote the space of modular "functions" of weight k over the congruence subgroup Γ we have the isomorphism of vector spaces

$$A_{2k}(\Gamma) \cong \Omega^k(X(\Gamma))$$

Hence, if we restrict ourselves to looking at modular "forms" (which are holomorphic, not only meromorphic) we have that $A_2k(\Gamma) = S_2(\Gamma)$ and so

$$S_2(\Gamma) \cong \Omega^1_{\text{hol}}(X(\Gamma))$$

Skipping over the actual construction of this isomorphism, we can see that there is a link between the forms on our modular curves and modular forms.

3. Now, we consider the Hecke operators T_n and the diamond operator $< \dot{>}$ (which takes $a \in (\mathbb{Z}/N\mathbb{Z})^\times$ and acts by the action of double coset $\Gamma\alpha\Gamma$ which is an element of the "Hecke Ring"). And so we have (without getting into details) that

- First, these operators act on the divisor groups of the modular curves $Div(X(\Gamma))$.
- Second, they act on the Jacobian, $Jac(X(\Gamma))$. Hence, by a theorem of Abel-Jacobi we know that they also act on $Pic^0(X(\Gamma)) \cong Jac(X(\Gamma))$, where the latter denotes the Picard group.
- Then, we find that the operators act on spaces of modular forms and it turns out that they are compatible with those on the jacobians of modular curves and relate through composition due to the fact that.

Continuing on-wards, an interesting (but not particularly well motivated) fact regarding this discussing is that we consider the latter objects such as $Jac(X(\Gamma))$ and $Div(X(\Gamma))$ as R-modules of the "Hecke Ring" as well as the fact that (in a particular case)

$$S_2(\Gamma_1(N))^*/\Lambda \cong Jac(X_1(N))$$

as $R-$modules where Λ is some nicely chosen subgroup. Moreover, we can consider the "Hecke algebra" generated by the operators

$$T_{\mathbb{Z}} = \mathbb{Z}[\{T_n, < n >: n \geqslant 0\}]$$

as well as the complex Hecke Algebra $T_{\mathbb{C}}$ which is the complex algebra generated by $T_{\mathbb{Z}}$ in $End(S_k(\Gamma))$. This leads us to prove one very nice result which states that

Proposition 3.2. *The space $S_k(\Gamma_0(N))^{new}$ of newforms admits a basis of eigenforms. The latter are forms which are simultaneous eigenvectors to all the Hecke operators T_n.*

Finally, for normalized eigenforms $f \in M_k(\Gamma_1(N))$, they admit an $L-$function which can be writen as

$$L(f, s) = \prod_p L_p(f, s)$$

where the product is taken over primes $p \nmid N$ and the local factor

$$L_p(f, s) = \sum_{n \geqslant 0} a_{p^n}(f) p^{-ns}$$

4. The next big step in the construction is to associate abelian varieties to modular forms $f \in S_2(\Gamma_0(N))^{new}$. We do this by considering the morphism

$$\lambda_f : T_{\mathbb{Z}} \to \mathbb{C}$$

such that $T_n(f) = \lambda_f(T)f \ \forall T_n \in T_{\mathbb{Z}}$. Letting $X_0(N) = X(\Gamma_0(N))$, we then recover an abelian variety A_f such that

$$A_f = Jac(X_0(N))/(ker(\lambda_f)Jac(X_0(N))$$

Heuristically, this is an abelian variety since $(ker(\lambda_f)Jac(X_0(N))$ is a subvariety of $Jac(X_0(N)$. Moreover, after another technical portion of the proof, we are lead to see that A_f can be defined as a variety over $\mathbb{Q}$.

5. We now introduce the "Eichler-Shimura" congruence relation. Essentially, it tell us that the following diagram commutes

$$\begin{array}{ccc} Div^0(X_0(N)) & \xrightarrow{T_p} & Div^0(X_0(N)) \\ \downarrow & & \downarrow \\ Pic^0(X_0(N)_p) & \xrightarrow{(Frob_p)_* + (Frob_p)^*} & Pic^0(X_0(N)_p) \end{array}$$

Moreover, after some quite involved arguments we also find that the diagram below commutes:

$$\begin{array}{ccc} Pic^0(X_0(N)) & \xrightarrow{T_p} & Pic^0(X_0(N)) \\ \downarrow & & \downarrow \\ Pic^0(X_0(N)_p) & \xrightarrow{(Frob_p)_* + (Frob_p)^*} & Pic^0(X_0(N)_p) \end{array}$$

Lets break this down:

- First, we find that at primes of good reduction (all but finitely many), the reduction mod p of the modular curve $X_0(N)_p$ remains an algebraic curve.
- The top relation is simply the Hecke operator discussed previously. The bottom relation is the pushforward + the pullback of the frobenius morphism between the picard groups of the reduced curves.
- By commutativity of the diagram we see that there is a relationship between the Hecke operator and the Frobenius morphism + its transpose.
- Finally, referring to the previous discussion (in 3) where we stated that an eigenform is a simultaneous eigenvectors to all the Hecke operators over a space, we will see that this allows us to recover information about its L-function and that of its associated Abelian variety (as in 4).

6. To bring everything together - let $f \in S_k(\Gamma_0(N))^{\text{new}}$ be some newform. Let A_f be its associated abelian variety constructed in point 4 (say $\dim(A_f) = d$). We are able to transfer the analytic maps from the 2nd commutative diagram in point 5 to algebraic maps and obtain the following diagram which commutes:

$$\begin{array}{ccc} A_f & \xrightarrow{T_p = a_p(f)} & A_f \\ \downarrow & & \downarrow \\ (A_f)_p & \xrightarrow{Frob_p + \widehat{Frob_p}} & (A_f)_p \end{array}$$

This is done by considering the surjective morphism $\alpha : Jac(X_0(N)) \to (A_f)_{\mathbb{C}}$ which is induced by the morphism $\hat{\alpha} : Pic^0(X_0(N)) \to A_f$. And so, by reading the diagram we immediately have that the reduction of the hecke operator T_p, which is related to the coefficients $a_p(f)$ is actually equal to $Frob_p + \widehat{Frob_p}$. Hence

$$(T_p)_p = Frob_p + \widehat{Frob_p} \in End((A_f)_p)$$

Then, by taking some ℓ-adic representation $\rho_p : End_{\mathbb{Q}}((A_f)_p) \to M_{2d}(\mathbb{Q}_\ell)$ for $\ell \neq p$ and a representation $\rho_{\mathbb{C}} : End(A_f) \to End(V_f^*)$ we chose a basis for $T_\ell(A_f)_p$ and from the Eichler-Shimura relation above we are able to compute the characteristic polynomial of $\rho_p(Frob_p)$ in terms of that of $\rho_{\mathbb{C}}(T_p)$. Now, recalling K_f, which is the feild $\mathbb{Q}$ adjoined by all of the coefficients of f - it turns out that $\rho_{\mathbb{C}}(T_p)$ is diagonal with respect to K_f-conjugates of f. The entries of $\rho_{\mathbb{C}}(T_p)$ are exactly the coefficients (eigenvalues of the operators) $a_p(f_\sigma)$. And so, without any explicit computations, we have an intuitive view as to why the $L-$function of A_f which involves the term $\rho_p(Frob_p)$ is related to the local factors $a_p(f_\sigma)$ of the $L-$function of f (and hence the global $L-$function of f).

To conclude the discussion regarding this theorem, I think it is interesting to note that when the dimension of the abelian variety A_f is 1 (i.e. it is an elliptic curve) we have a theorem which gives us a relationship between L-functions of modular forms and elliptic curves. In the years following the paper of Shimura, it is the converse of the theorem above which led to the famous Taniyama–Shimura-Weil Conjecture (now known as the Modularity Theorem) stating that to each elliptic curve E defined over $\mathbb{Q}$ there exists a modular form f which is related to it by its L-function, i.e. $L(f, s) = L(E/\mathbb{Q}, s)$.

References

[1] J. H. Silverman, *The Arithmetic of Elliptic Curves*, Second Edition, New York: Springer-Verlag, 2009.
[2] Goro Shimura, *Introduction to the arithmetic theory of automorphic functions*, Publ. of Math. Soc. of Japan, 11, 1971
[3] Neal Koblitz, *Introduction To Elliptic Curves And Modular Forms*, Springer-Verlag 1984
[4] https://corentinperretgentil.gitlab.io/static/documents/eichler-shimura.pdf
[5] http://homepages.warwick.ac.uk/ mariaz/eichshim3.pdf

Department of Mathematics and Statistics, McGill University
Email address: andreas.hatziiliou@mail.mcgill.ca