25 Modular forms and L-series

As we will show in the next lecture, Fermat's Last Theorem is a direct consequence of the following theorem [11, 12].

Theorem 25.1 (Taylor-Wiles)**.** *Every semistable elliptic curve $E/\mathbb{Q}$ is modular.*

In fact, as a result of subsequent work [3], we now have the stronger result, proving what was previously known as the modularity conjecture (or Taniyama-Shimura-Weil conjecture).

Theorem 25.2 (Breuil-Conrad-Diamond-Taylor)**.** *Every elliptic curve $E/\mathbb{Q}$ is modular.*

Our goal in this lecture is to explain what it means for an elliptic curve over $\mathbb{Q}$ to be modular (we will also define the term semistable). This requires us to delve briefly into the theory of modular forms. Our goal in doing so is simply to understand the definitions and the terminology; we will omit all but the most straight-forward proofs.

25.1 Modular forms

Definition 25.3. A holomorphic function $f\colon \mathbb{H} \to \mathbb{C}$ is a *weak modular form* of *weight* k for a congruence subgroup Γ if

$$f(\gamma\tau) = (c\tau + d)^k f(\tau)$$

for all $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma$.

The j-function $j(\tau)$ is a weak modular form of weight 0 for $\mathrm{SL}_2(\mathbb{Z})$, and $j(N\tau)$ is a weak modular form of weight 0 for $\Gamma_0(N)$. For an example of a weak modular form of positive weight, recall the Eisenstein series

$$G_k(\tau) := G_k([1,\tau]) := \sum_{\omega\in[1,\tau]}{}' \frac{1}{\omega^k} = \sum_{m,n\in\mathbb{Z}}{}' \frac{1}{(m+n\tau)^k},$$

which, for $k \geq 3$, is a weak modular form of weight k for $\mathrm{SL}_2(\mathbb{Z})$.[1] To see this, recall that $\mathrm{SL}_2(\mathbb{Z})$ is generated by the matrices $S = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ and $T = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$, and note that

$$G_k(S\tau) = G_k(-1/\tau) = \sum_{m,n\in\mathbb{Z}}{}' \frac{1}{(m-\frac{n}{\tau})^k} = \sum_{m,n\in\mathbb{Z}}{}' \frac{\tau^k}{(m\tau-n)^k} = \tau^k G_k(\tau),$$
$$G_k(T\tau) = G_k(\tau+1) = G_k(\tau) = 1^k G(\tau).$$

If Γ contains $-I$, than any weakly modular function f for Γ must satisfy $f(\tau) = (-1)^k f(\tau)$, since $-I$ acts trivially but $c\tau + d = -1$; in this case only the zero function can be a weak modular form of odd weight for Γ. We are specifically interested in the congruence subgroup $\Gamma_0(N)$, which contains $-I$, so we will restrict our attention to modular forms of even weight.[2]

[1] Recall that the notation $\sum'$ means that terms whose denominators vanish are to be omitted.

[2] Some authors use $2k$ in place of k for this reason, but we should note that for other congruence subgroups such as $\Gamma_1(N)$ that do not contains -1 (for $N > 2$) there are interesting modular forms of odd weight.

As we saw with modular functions (see Lecture 20), if Γ is a congruence subgroup of level N, meaning that it contains $\Gamma(N)$, then Γ contains the matrix $T^N = \left(\begin{smallmatrix} 1 & N \\ 0 & 1 \end{smallmatrix}\right)$, and every weak modular form $f(\tau)$ for Γ must satisfy $f(\tau+N) = f(\tau)$ for $\tau \in \mathbb{H}$, since for the matrix T^N we have $c = 0$ and $d = 1$, so $(c\tau + d)^k = 1^k = 1$. It follows that $f(\tau)$ has a q-expansion of the form

$$f(\tau) = f^*(q^{1/N}) = \sum_{n=-\infty}^{\infty} a_n q^{n/N},$$

where $q = e^{2\pi i \tau}$. We say that f is *holomorphic at* ∞ if f^* is holomorphic at 0, equivalently, $a_n = 0$ for all $n < 0$. We say that f is *holomorphic at the cusps* if $f(\gamma\tau)$ is holomorphic at ∞ for all $\gamma \in \mathrm{SL}_2(\mathbb{Z})$. As with modular functions, we only need to check this condition at a finite set of cusp representatives for Γ.

Definition 25.4. A *modular form* f is a weak modular form that is holomorphic at the cusps. Equivalently, f is a weak modular form that extends to a holomorphic function on the extended upper half plane $\mathbb{H}^* = \mathbb{H} \cup \mathbb{P}^1(\mathbb{Q})$.

The only modular forms of weight 0 are constant functions. This is main motivation for introducing the notion of weight, it allows us to generalize modular functions in an interesting way, by strengthening their analytic properties (holomorphic on $\mathbb{H}^*$, not just meromorphic) at the expense of weakening their congruence properties (modular forms of positive weight are not Γ-invariant due to the factor $(c\tau + d)^k$).

The j-function is not a modular form, since it has a pole at ∞, but the Eisenstein functions $G_k(\tau)$ are modular forms of weight k for $\mathrm{SL}_2(\mathbb{Z})$ (as noted above, only even k are interesting). For $\Gamma = \mathrm{SL}_2(\mathbb{Z})$ there is just on cusp to check and it suffices to note that for even integer $k \geq 4$ we have

$$\lim_{\mathrm{im}\,\tau\to\infty} G_k(\tau) = \lim_{\mathrm{im}(\tau)\to\infty} \sum_{m,n\in\mathbb{Z}}{}' \frac{1}{(m+n\tau)^k} = 2\sum_{n=1}^{\infty} \frac{1}{n^k} = 2\zeta(k) < \infty,$$

(recall that the series converges absolutely, which justifies rearranging its terms).

Definition 25.5. A modular form is a *cusp form* if it vanishes at all the cusps. Equivalently, its q-expansion at every cusp has constant coefficient $a_0 = 0$

Example 25.6. For even k the Eisenstein series $G_k(\tau)$ is not a cusp form, but the discriminant function

$$\Delta(\tau) = g_2(\tau)^3 - 27g_3(\tau)^2,$$

with $g_2(\tau) = 60G_4(\tau)$ and $g_3(\tau) = 140G_6(\tau)$, is a cusp form of weight 12 for $\mathrm{SL}_2(\mathbb{Z})$; to see that it vanishes ∞, note that $j(\tau) = g_2(\tau)^3/\Delta(\tau)$ has a pole at ∞ and $g_2(\tau)$ does not, so $\Delta(\tau)$ must vanish (see the proof of Theorem 16.11).

If f and g are two modular forms of weight k for Γ, then their sum $f + g$ is also a modular form of weight k for γ, as is λf for any $\lambda \in \mathbb{C}$. Thus for any fixed weight k and congruence subgroup Γ we obtain a $\mathbb{C}$-vector space $M_k(\Gamma)$ that contains the cusp forms $S_k(\Gamma)$ as a subspace.[3] The dimension of these vector spaces is finite, and can be explicitly in terms of invariants of the corresponding modular curve $X(\Gamma) = \mathbb{H}^*/\Gamma$.

[3]Products of modular forms are also modular forms (of higher weight), but we will not use this.

As in Problem Set 10, let $\nu_2(\Gamma)$ count the number of Γ-inequivalent $\mathrm{SL}_2(\mathbb{Z})$-translates of i fixed by some $\gamma \in \Gamma$ other than $\pm I$ (elliptic points of period 2), and similarly define $\nu_3(\Gamma)$ in terms of $\rho = e^{2\pi i/3}$ (elliptic points of period 3). Let ν_∞ denote the number of cusp orbits in $X(\Gamma)$, and $g(\Gamma)$ the genus of $X(\Gamma)$.

Theorem 25.7. *Let Γ be a congruence subgroup. For $k = 0$ we have $\dim M_k(\Gamma) = 1$ and $\dim S_k(\Gamma) = 0$. For any even integer $k > 0$ we have*

$$\dim M_k(\Gamma) = (k-1)(g(\Gamma)-1) + \left\lfloor \frac{k}{4} \right\rfloor \nu_2 + \left\lfloor \frac{k}{3} \right\rfloor \nu_3 + \frac{k}{2}\nu_\infty,$$

and if $k > 2$ we also have

$$\dim S_k(\Gamma) = (k-1)(g(\Gamma)-1) + \left\lfloor \frac{k}{4} \right\rfloor \nu_2 + \left\lfloor \frac{k}{3} \right\rfloor \nu_3 + \left(\frac{k}{2} - 1\right)\nu_\infty.$$

For $k = 2$ we have $\dim S_k(\Gamma) = g(\Gamma)$.

Proof. See [4, Thm. 3.5.1] □

We are specifically interested in the vector space $S_2(\Gamma_0(N))$ of dimension $g(\Gamma_0(N))$.

Remark 25.8. Those who know a bit of algebraic geometry may have guessed that there is a relationship between cusp forms in $S_2(\Gamma_0(N))$ and the space of regular differentials for the modular curve $X_0(N)$, since their dimensions coincide; this is indeed the case.

25.2 Hecke operators

In order to understand the relationship between modular forms and elliptic curves we need to construct a suitable basis for $S_2(\Gamma_0(N))$. To help with this, we now introduce the *Hecke operators*. For each positive integer n, the Hecke operator T_n is a linear operator that we can apply to any of the vector spaces $M_k(\Gamma_0(N))$, and it fixes the subspace of cusp forms, so it is also a linear operator on $S_k(\Gamma_0(N))$.[4]

In order to motivate the definition of the Hecke operators on modular forms, we first define them in terms of lattices, following the presentation in [7, VII.5.1]. As in previous lectures, a lattice (in $\mathbb{C}$) is an additive subgroup of $\mathbb{C}$ that is a free $\mathbb{Z}$-module of rank 2 containing an $\mathbb{R}$-basis for $\mathbb{C}$.

For each positive integer n, the Hecke operator T_n sends each lattice $L = [\omega_1, \omega_2]$ to the formal sum of its index-n sublattices.

$$T_n L := \sum_{[L:L']=n} L'. \tag{1}$$

Here we are working in the free abelian group $\mathrm{Div}(\mathcal{L})$ generated by the set $\mathcal{L}$ of all lattices; we extend T_n linearly to an endomorphism of $\mathrm{End}(\mathrm{Div}(\mathcal{L}))$ (this means $T_n \sum L := \sum T_n L$). Another family of endomorphisms of $\mathrm{Div}(L)$ are the homethety operators R_λ defined by

$$R_\lambda L := \lambda L, \tag{2}$$

for any $\lambda \in \mathbb{C}^\times$. This setup might seem overly abstract, but it allows one to easily prove some essential properties of the Hecke operators that are applicable in many settings. When defined in this generality the Hecke operators are also sometimes called *correspondences*.

[4] One can define Hecke operators more generally on $M_k(\Gamma_1(N))$ (which contains $M_k(\Gamma_0(N))$), but the definition is more involved and not needed here.

Remark 25.9. Recall that if $E/\mathbb{C}$ is the elliptic curve isomorphic to the torus $\mathbb{C}/L$, the index-n sublattices of L correspond to n-isogenous elliptic curves. The fact that the Hecke operators average over sublattices is related to the fact that the relationship between modular forms and elliptic curves occurs at the level of isogeny classes.

Theorem 25.10. *The operators T_n and R_λ satisfy the following:*

(i) $T_n R_\lambda = R_\lambda T_n$ *and* $R_\lambda R_\mu = R_{\lambda\mu}$.

(ii) $T_{mn} = T_m T_n$ *for all* $m \perp n$.

(iii) $T_{p^{r+1}} = T_{p^r} T_p - p T_{p^{r-1}} R_p$ *for all primes p and integers $r \geq 1$.*

Proof. (i) is clear, as is (ii) if we note that for $m \perp n$ there is a bijection between index mn-sublattices L'' of L and pairs (L', L'') with $[L : L'] = n$ and $[L' : L''] = m$. For (iii), the first term on the RHS counts pairs (L', L'') with $[L : L'] = p$ and $[L' : L''] = p^r$, and the second term corrects for over counting; see [7, Prop. VII.10] for the details. □

Corollary 25.11. *The sub-ring of* $\mathrm{End}(\mathrm{Div}(\mathcal{L}))$ *generated by* $\{R_p, T_p : p \text{ prime}\}$ *is commutative and contains all the T_n.*

Proof. By recursively applying (iii) we can reduce any T_{p^r} to a polynomial in T_p and R_p, and any two such polynomials commute (since T_p and R_p commute, by (i)). Moreover, (i) and (ii) imply that for distinct primes p and q, polynomials in T_p, R_p commute with polynomials in T_q, R_q. Using (ii) and (iii) we can reduce any T_n to a product of polynomials in T_{p_i}, R_{p_i} for distinct primes p_i and the corollary follows. □

Any function $F\colon \mathcal{L} \to \mathbb{C}$ extends linearly to a function $F\colon \mathrm{Div}(\mathcal{L}) \to \mathbb{C}$ to which we may apply any operator $T \in \mathrm{End}(\mathrm{Div}(\mathcal{L}))$, yielding a new function $TF\colon \mathrm{Div}(\mathcal{L}) \to \mathbb{C}$ defined by $TF\colon D \mapsto F(T(D)$; restricting TF to $\mathcal{L} \subseteq \mathrm{Div}(\mathcal{L})$ then gives a function $TF\colon \mathcal{L} \to \mathbb{C}$ that we regard as the transform of our original function F by T. This allows us to apply the Hecke operators T_n and homethety operators R_λ to any function that maps lattices to complex numbers. We work this out explicitly for the Hecke operators acting on modular forms for $\mathrm{SL}_2(\mathbb{Z})$ in the next section.

25.3 Hecke operators for modular forms of level one

We now define the action of the Hecke operators T_n on $M_k(\mathrm{SL}_2(\mathbb{Z}) = M_k(\Gamma_0(1))$. The case $M_k(\Gamma_0(N))$ is analogous, but the details are more involved, so let us fix $N = 1$ for the purposes of presentation and address $N > 1$ in the remarks.

Let f be a modular form of weight k. We can view any view $f(\tau)$ as a function on lattices $[1, \tau]$, which we extend to arbitrary lattices $L = [\omega_1, \omega_2]$ by defining

$$f([\omega_1, \omega_2]) := f(\omega_1^{-1}[1, \omega_2/\omega_1]) = \omega_1^{-k} f([1, \omega_2/\omega_1]),$$

we assume ω_1 and ω_2 are ordered so that ω_2/ω_1 is in the upper half plane. Conversely, any function $F\colon \mathcal{L} \to \mathbb{C}$ on lattices induces a function $\tau \mapsto F([1, \tau])$ on the upper half plane. Viewing our modular form f as a function $\mathcal{L} \to \mathbb{C}$, we can transform this function by any $T \in \mathrm{End}(\mathrm{Div}(\mathcal{L}))$ as described above, thereby obtaining a new function $\mathcal{L} \to \mathbb{C}$ that induces a function $Tf\colon \mathbb{H} \to \mathbb{C}$ on the upper half plane.

For any $f \in M_k(\Gamma_0(1))$ we thus define $R_\lambda f$ as

$$R_\lambda f(\tau) := f(\lambda[1, \tau]) = \lambda^{-k} f(\tau),$$

which clearly lies in $M_k(\Gamma_0(1))$, and if f is a cusp form, so is $R_\lambda f$.

We define $T_n f$ similarly, but introduce a scaling factor of n^{k-1} that simplifies the formulas that follow. An easy generalization of Lemma 21.2 shows that for each integer $n \geq 1$, the index n sublattices of $[1, \tau]$ are given by

$$\Big\{[d, a\tau + b] : ad = n,\ 0 \leq b < d\Big\};$$

see [7, Lemma VII.5.2], for example. If we rescale by d^{-1} to put them in the form $[1, \omega]$, we have $\omega = (a\tau + b)/d$. For $f \in M_k(\Gamma_0(1))$ we thus define $T_n f$ as

$$T_n f(\tau) := n^{k-1} \sum_{[L:L']=n} f(L) = n^{k-1} \sum_{ad=n,\, 0\leq b<d} d^{-k} f\left(\frac{a\tau + b}{d}\right),$$

which is also clearly an element of $M_k(\Gamma_0(1))$, and if f is a cusp form, so is $T_n f$. It is clear from the definition that T_n acts linearly, so it is a linear operator on the vector spaces $M_k(\Gamma_0(1))$ and $S_k(\Gamma_0(1))$. Theorem 25.10 then yields the following corollary.

Corollary 25.12. *The Hecke operators T_n for $M_k(\Gamma_0(1))$ satisfy $T_{mn} = T_m T_n$ for $m \perp n$ and $T_{p^{r+1}} = T_{p^r} T_p - p^{k-1} T_{p^{r-1}}$ for p prime.*

Proof. The first equality is clear; the second term on the RHS of the second equality arises from the fact that $pT_{p^{r-1}} R_p f = p^{k-1} T_{p^{r-1}} f$. □

The corollary implies that it suffices to understand the behavior of T_p for p prime. Let us compute the the q-series expansion of $T_p f$, where $f(\tau) = \sum_{n=1}^\infty a_n q^n$ is a cusp form of weight k for $\Gamma_0(1)$. We have

$$\begin{aligned}
T_p f(\tau) &= p^{k-1} \sum_{\substack{ad=p \\ 0\leq b<d}} d^{-k} f\left(\frac{a\tau + b}{d}\right) \\
&= p^{k-1} f(p\tau) + p^{-1} \sum_{b=0}^{p-1} f\left(\frac{\tau + b}{p}\right) \\
&= p^{k-1} \sum_{n=1}^\infty a_n e^{2\pi i n p} + p^{-1} \sum_{b=0}^{p-1} \sum_{n=1}^\infty a_n e^{2\pi i n(\tau+b)/p} \\
&= p^{k-1} \sum_{n=1}^\infty a_n q^{np} + p^{-1} \sum_{b=0}^{p-1} \sum_{n=1}^\infty a_n \zeta_p^{bn} q^{n/p} \\
&= p^{k-1} \sum_{n=1}^\infty a_{n/p} q^n + p^{-1} \sum_{n=1}^\infty a_n \left(\sum_{b=0}^{p-1} \zeta_p^{bn}\right) q^{n/p} \\
&= \sum_{n=1}^\infty \left(a_{np} + p^{k-1} a_{n/p}\right) q^n,
\end{aligned}$$

where $\zeta_p = e^{2\pi i/p}$ and we define $a_{n/p}$ to be 0 if p does not divide n. This calculation yields the following theorem and corollary, in which we use $a_n(f)$ to denote the coefficient of q^n in the q-expansion of f, for any modular form f and integer n.

Theorem 25.13. *For any $f \in S_k(\Gamma_0(1))$ and prime p we have*

$$a_n(T_p f) = \begin{cases} a_{np}(f) & \text{if } p \nmid n, \\ a_{np}(f) + p^{k-1} a_{n/p}(f) & \text{if } p \mid n. \end{cases}$$

Corollary 25.14. *For any modular form $f \in S_k(\Gamma_0(1))$ and relatively prime integers m and n we have $a_m(T_n f) = a_{mn}(f)$. In particular, $a_1(T_n f) = a_n(f)$.*

Proof. The corollary follows immediately from Theorem 25.13 for n prime. For composite n (and any $m \perp n$), we proceed by induction on n. If $n = cd$ with $c \perp d$ both greater than 1, then by Theorem 25.13 and the inductive hypothesis we have

$$a_m(T_n f) = a_m(T_c T_d f) = a_{mc}(T_d f) = a_{mcd} = a_{mn}.$$

For $n = p^{r+1}$, applying Theorem 25.13, Corollary 25.12, and the inductive hypothesis yields

$$\begin{aligned} a_m(T_{p^{r+1}} f) &= a_m(T_{p^r} T_p f) - p^{k-1} a_m(T_{p^{r-1}} f) \\ &= a_{mp^r}(T_p f) - p^{k-1} a_{mp^{r-1}}(f) \\ &= a_{mp^{r+1}}(f) + p^{k-1} a_{mp^{r-1}}(f) - p^{k-1} a_{mp^{r-1}}(f) \\ &= a_{mn}(f), \end{aligned}$$

as desired. □

Remark 25.15. All the results in this section hold for $f \in S_k(\Gamma_0(N))$ if we restrict to Hecke operators T_n with $n \perp N$, which is all that we require, and the key result $a_1(T_n f) = a_n(f)$ holds in general. For $p|N$ the definition of T_p (and T_n for $p|n$) needs to change and the formulas in Corollary 25.12 and Theorem 25.13 must be modified. The definition of the Hecke operators is more complicated (in particular, it depends on the level N), but some of the formulas are actually simpler (for example, for $p|N$ we have $T_{p^r} = T_p^r$).

25.4 Eigenforms

The Hecke operators T_n form an infinite family of linear operators on the vector space $S_k(\Gamma_0(1))$. We are interested in the elements $f \in S_k(\Gamma_0(1))$ that are simultaneous eigenvectors for all of the Hecke operators; this means that for every $n \geq 1$ we have $T_n f = \lambda_n f$ for some eigenvalue $\lambda_n \in \mathbb{C}^*$ of T_n. When such an f also satisfies $a_1(f) = 1$, we call it a (normalized) *eigenform*. We then we know exactly what the Hecke eigenvalues λ_n of an eigenform $f = \sum a_n q^n$ must be: if $T_n f = \lambda_n f$ then we have

$$\lambda_n a_1 = a_1(T_n f) = a_n(f) = a_n,$$

by Corollary 25.14, and $a_1 = 1$ so $\lambda_n = a_n$. Thus the Hecke eigenvalue λ_n are precisely the coefficients a_n in the q-expansion of f. Corollary 25.12 implies that the a_n must then satisfy

$$\begin{aligned} a_{mn} &= a_m a_n && (m \perp n) \\ a_{p^r} &= a_p a_{p^{r-1}} - p^{k-1} a_{p^{r-2}} && (p \text{ prime}) \end{aligned}$$

For $k = 2$ the second recurrence above should look familiar — it is exactly the same as the recurrence satisfied by the Frobenius traces $a_{p^r} = p^r + 1 - \#E(\mathbb{F}_{p^r})$ of an elliptic curve $E/\mathbb{F}_p$, as you proved in Problem Set 7.

Our goal is to construct a basis of eigenforms for $S_k(\Gamma_0(1))$, and to prove that it is unique. In order to do so, we need to introduce the *Petersson inner product*. Recall that if $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_0(1)$, then $\operatorname{im}\gamma\tau = \operatorname{im}\tau/|c\tau+d|^2$, thus for $f, g \in S_k(\Gamma_0(1))$ we have

$$f(\gamma\tau)\overline{g(\gamma\tau)}(\operatorname{im}\gamma\tau)^k = (c\tau+d)^k f(\tau)(c\bar{\tau}+d)^k g(\tau)\left(\frac{\operatorname{im}\tau}{|c\tau+d|^2}\right)^k = f(\tau)\overline{g(\tau)}(\operatorname{im}\tau)^k.$$

The function $f(\tau)\overline{g(\tau)}(\operatorname{im}\tau)^k$ is thus $\Gamma_0(1)$-invariant. If we parameterize the upper half-plane $\mathbb{H}$ with real parameters $x = \operatorname{re}\tau$ and $y = \operatorname{im}\tau$, so $\tau = x + iy$, it is straight-forward to check that the measure

$$\mu(U) = \iint_U \frac{dxdy}{y^2}$$

is $\Gamma_0(1)$-invariant, that is, $\mu(\gamma U) = \mu(U)$ for all measurable sets $U \subseteq \mathbb{H}$. This motivates the following definition.

Definition 25.16. The *Petersson inner product* on $S_k(\Gamma_0(1))$ is defined by

$$\langle f, g\rangle = \int_{\mathcal{F}} f(\tau)\overline{g(\tau)}y^{k-2}dxdy, \tag{3}$$

where the integral ranges over points $\tau = x + yi$ in a fundamental region $\mathcal{F}$ for $\Gamma_0(1)$.

It is easy to check that $\langle f, g\rangle$ is a positive definite Hermitian form: it is bilinear in both f and g, it satisfies $\langle f, g\rangle = \overline{\langle g, f\rangle}$, and $\langle f, f\rangle \geq 0$ with equality only when $f = 0$. Thus it defines an inner product on the complex vector space $S_k(\Gamma_0(1))$.

One can show that the Hecke operators are self-adjoint with respect to the Petersson inner product, that is, they satisfy $\langle f, T_n g\rangle = \langle T_n f, g\rangle$. The T_n are thus Hermitian (normal) operators, and we know from Corollary 25.12 that they all commute with each other. This makes it possible to apply the following form of the Spectral Theorem.

Lemma 25.17. *Let V be a finite-dimensional $\mathbb{C}$-vector space equipped with a positive definite Hermitian form, and let $\alpha_1, \alpha_2, \ldots$ be a sequence of commuting Hermitian operators. Then $V = \bigoplus_i V_i$, where each V_i is an eigenspace of every α_n.*

Proof. The matrix for α_1 is Hermitian, therefore diagonalizable,[5], so we can decompose V as a direct sum of eigenspaces for α_1, writing $V = \bigoplus_i V(\lambda_i)$, where the λ_i are the distinct eigenvalues of α_1. Because α_1 and α_2 commute, α_2 must fix each subspace $V(\lambda_i)$, since for each $v \in V(\lambda_i)$ we have $\alpha_1\alpha_2 v = \alpha_2\alpha_1 v = \alpha_2\lambda_i v = \lambda_i\alpha_2 v$, and therefore $\alpha_2 v$ is an eigenvector for α_1 with eigenvalue λ_i, so $\alpha_2 v \in V(\lambda_i)$. Thus we can decompose each $V(\lambda_i)$ as a direct sum of eigenspaces for α_2, and may continue in this fashion for all the α_n. □

By Lemma 25.17, we may decompose $S_k(\Gamma_0(1)) = \bigoplus_i V_i$ as a direct sum of eigenspaces for the Hecke operators T_n. Let $f(\tau) = \sum a_n q^n$ be a nonzero element of V_i. We $a_1(T_n f) = a_n$, by Corollary 25.12, and also $T_n f = \lambda_n f$, for some eigenvalue λ_n of T_n which is determined by V_i, so $a_n = \lambda_n a_1$. This implies $a_1 \neq 0$, since otherwise $f = 0$, and if we normalize f so that $a_1 = 1$ (which we can do, since f is nonzero and V_i is a $\mathbb{C}$-vector space), we then have $a_n = \lambda_n$ for all $n \geq 1$, and f completely determined by the sequence of Hecke eigenvalues λ_n for V_i. It follows that every element of V_i is a multiple of f, so $\dim V_i = 1$ and the eigenforms in $S_k(\Gamma_0(1))$ form a basis.

[5]This fact is also sometimes called the Spectral Theorem and proved in most linear algebra courses.

Theorem 25.18. *The vector space $S_k(\Gamma_0(1))$ can be written as a direct sum of one-dimensional eigenspaces for the Hecke operators T_n and has a unique basis of eigenforms $f(\tau) = \sum a_n q^n$, where each a_n is the eigenvalue of T_n on the 1-dimensional subspace generated by f.*

Remark 25.19. Theorem 25.18 fails for $S_k(\Gamma_0(N))$ for two reasons, both of which are readily addressed. First, as in 25.15, we need to restrict our attention to the Hecke operators T_n with $n \perp N$ (when n and N have a common factor T_n is not necessarily a Hermitian operator with respect to the Petersson inner product). We can then proceed as above to decompose $S_k(\Gamma_0(N))$ into subspaces whose elements are simultaneous eigenvectors for all the T_n with $n \perp N$, but these subspaces need not be one-dimensional. In order to ensure this we restrict our attention to a particular subspace of $S_k(\Gamma_0(N))$. We call a cusp form $f \in S_k(\Gamma_0(N))$ an *oldform* if it also lies in $S_k(\Gamma_0(M))$ for some $M|N$, which we note is a subspace of $S_k(\Gamma_0(N))$ (since $\Gamma_0(M)$-invariance implies $\Gamma_0(N)$-invariance for $M|N$). The oldforms in $S_k(\Gamma_0(N))$ generate a subspace $S_k^{\text{old}}(\Gamma_0(N))$, and we define $S_k^{\text{new}}(\Gamma_0(N))$ as the orthogonal complement of $S_k^{\text{old}}(\Gamma_0(N))$ under the Petersson inner product, so that

$$S_k(\Gamma_0(N)) = S_k^{\text{old}}(\Gamma_0(N)) \oplus S_k^{\text{new}}(\Gamma_0(N)),$$

and we call the eigenforms in $S_k^{\text{new}}(\Gamma_0(N))$ *newforms*. One can show that the Hecke operators T_n with $n \perp N$ preserve both $S_k^{\text{old}}(\Gamma_0(N))$ and $S_k^{\text{new}}(\Gamma_0(N))$. If we then decompose $S_k^{\text{new}}(\Gamma_0(N))$ into eigenspaces with respect to these operators, the resulting eigenspaces are all one-dimensional, moreover, each is actually generated by an eigenform (a simultaneous eigenvector for *all* the T_n, not just those with $n \perp N$ that we used to obtain the decomposition); this is a famous result of Atkin and Lehner [2, Thm. 5]. Thus Theorem 25.18 remains true if we simply replace $S_k(\Gamma_0(1))$ by $S_k^{\text{new}}(\Gamma_0(N))$.

25.5 The *L*-series of a modular form

Our interest in cusp forms is each has an associated *L-series*.

Definition 25.20. The *L-series* of a cusp form $f(\tau) = \sum_{n=1}^{\infty} a_n q^n$ is the function

$$L_f(s) = \sum_{n=1}^{\infty} a_n n^{-s}.$$

The series for $L_f(s)$ converges uniformly for $\text{re}(s) > 1 + k/2$, where k is the weight of f; it is an example of a *Dirichlet series*, a complex function of the form $f(x) = \sum_{n=1}^{\infty} a_n n^{-2}$.

Example 25.21. The most famous Dirichlet series is the *Riemann zeta function*

$$\zeta(s) = \sum_{n=1}^{\infty} n^{-s}.$$

It converges only for $\text{re}(s) > 1$, but it has an analytic continuation that is holomorphic everywhere except at $s = 1$, where it has a simple pole. The *normalized* zeta function[6]

$$\tilde{\zeta}(s) = \pi^{-s/2}\Gamma(s/2)\zeta(s),$$

[6]Here $\Gamma(s) = \int_0^{\infty} e^{-t}t^{s-1}dt$ denotes the gamma function.

satisfies the *functional equation*

$$\tilde{\zeta}(s) = \tilde{\zeta}(1-s).$$

The Riemann zeta function can also be written as an *Euler product*

$$\zeta(s) = \prod_p (1-p^{-s})^{-1} = \prod_p (1+p^{-s}+p^{-2s}+\dots) = \sum_{n=1}^{\infty} n^{-s}.$$

The L-series $L_f(s)$ of a cusp form f for $\Gamma_0(1)$ also has an analytic continuation, a functional equation.

Theorem 25.22 (Hecke)**.** *Let $L_f(s)$ be the L-series of $f \in S_k(\Gamma_0(N))$. Then $L_f(s)$ extends analytically to a holomorphic function on $\mathbb{C}$, and the normalized function*

$$\tilde{L}_f(s) = N^{s/2}(2\pi)^{-s}\Gamma(s)L_f(s).$$

satisfies a functional equation of the form

$$\tilde{L}_f(s) = \pm\tilde{L}_f(k-s).$$

Remark 25.23. There are more explicit version of this theorem that determine the sign in the functional equation above.

In the case that $f \in S_k(\Gamma_0(N))$ is an eigenform we also get an Euler product for $L_f(s)$; as noted in Remark 25.19, this means we want to restrict to newforms.

Theorem 25.24. *Let $L_f(s)$ be the L-series of an eigenform $f \in S_k^{\text{new}}(\Gamma_0(N))$. Then $L_f(s)$ has the Euler product*

$$L_f(s) = \sum_{n=1}^{\infty} a_n n^{-s} = \prod_p (1 - a_p p^{-s} + \chi(p)p^{k-1}p^{-2s})^{-1}, \tag{4}$$

where $\chi(p) = 0$ for $p|N$ and $\chi(p) = 1$ otherwise.

25.6 The *L*-series of an elliptic curve

What does all this have to do with elliptic curves? Like eigenforms, elliptic curves over $\mathbb{Q}$ also have an L-series with an Euler product. In fact, with elliptic curves, we use the Euler product to define the L-series.

Definition 25.25. The *L-series of an elliptic curve $E/\mathbb{Q}$* is given by the Euler product

$$L_E(s) = \prod_p L_p(p^{-s})^{-1} = \prod_p \left(1 - a_p p^{-s} + \chi(p)pp^{-2s}\right)^{-1}, \tag{5}$$

where $\chi(p)$ is 0 if E has bad reduction at p, and 1 otherwise.[7] For primes p where E has good reduction (all but finitely many), a_p is the Frobenius trace $p+1-\#E_p(\mathbb{F}_p)$, where E_p denotes the reduction of E modulo p. Equivalently, the polynomial $L_p(T)$ is the numerator of the zeta function

$$Z(E_p;T) = \exp\left(\sum_{n=1}^{\infty} \#E_p(\mathbb{F}_{p^n})\frac{T^n}{n}\right) = \frac{1-a_pT+T^2}{(1-T)(1-pT)},$$

[7] As explained in §25.7, this assumes we are using a minimal Weierstrass equation for E.

that appeared in Problem Set 7. For primes p where E has bad reduction, the polynomial $L_p(T)$ is defined by

$$L_p(T) = \begin{cases} 1 & \text{if } E \text{ has } \textit{additive} \text{ reduction at } p. \\ 1-T & \text{if } E \text{ has } \textit{split mulitiplicative} \text{ reduction at } p. \\ 1+T & \text{if } E \text{ has } \textit{non-split multiplicative} \text{ reduction at } p. \end{cases}$$

according to the type of reduction E has at p, as explained in the next section. This means that $a_p \in \{0, \pm 1\}$ at bad primes.

The L-series $L_E(s)$ converges for $\mathrm{re}(s) > 3/2$. As we will see shortly, the question of whether or not $L_E(s)$ has an analytic continuation is intimately related to the question of modularity (we now know the answer is yes, since every elliptic curve over $\mathbb{Q}$ is modular).

25.7 The reduction type of an elliptic curve

When computing $L_E(s)$, it is important to use a *minimal Weierstrass equation* for E, one that has good reduction at as many primes as possible. To see why this is necessary, note that if $y^2 = x^3 + Ax + B$ is a Weierstrass equation for E, then, up to isomorphism, so is $y^2 + u^4Ax + u^6B$, for any integer u, and this equation will have bad reduction at all primes $p|u$. Moreover, even though the equation $y^2 = x^3 + Ax + B$ always has bad reduction at 2, there may be an isomorphic equation in general Weierstrass form that has good reduction at 2. For example, the elliptic curve defined by $y^2 = x^3 + 16$ is isomorphic to the elliptic curve defined by $y^2 + y = x^3$ (replace x by $4x$, divide by 64, and then replace y by $y + 1/2$), which does have good reduction at 2.

Definition 25.26. Let $E/\mathbb{Q}$ be an elliptic curve. A *minimal Weierstrass equation* for E is a general Weierstrass equation

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$$

with $a_1, a_2, a_3, a_4, a_6 \in \mathbb{Z}$ that defines an elliptic curve $E'/\mathbb{Q}$ that is isomorphic to E over $\mathbb{Q}$ and whose discriminant $\Delta(E')$ divides the discriminant of every elliptic curve isomorphic to E. The discriminant $\Delta(E')$ is the *minimal discriminant* of E and is denoted $\Delta_{\min}(E)$.

It is not immediately obvious that an elliptic curve necessarily has a minimal Weierstrass equation, but for elliptic curves over $\mathbb{Q}$ this is indeed the case; see [8, Prop. VII.1.3]. It can be computed in Sage via `E.minimal_model()`; see [5] for an explicit algorithm.

We now address the three types of bad reduction. To simplify matters, we are going to ignore the prime 2. At any odd prime p of bad reduction we can represent $E_p/\mathbb{F}_p$ by an equation of the form $y^2 = f(x)$, for some cubic $f \in \mathbb{F}_p[x]$ that has a repeated root. We can choose $f(x)$ so that this repeated root is at 0, and it is easy to verify that the projective curve defined by the equation $y^2 = f(x)$ has exactly one singular point, which occurs at the affine point $(0, 0)$.

If we exclude the point $(0, 0)$, the standard algebraic formulas for the group law on $E(\mathbb{F}_p)$ still work, and the set

$$E_p^{\mathrm{ns}}(\mathbb{F}_p) = E_p(\mathbb{F}_p) - \{(0, 0)\}$$

of non-singular points of $E_p(\mathbb{F}_p)$ is actually closed under the group operation. Thus $E_p^{\mathrm{ns}}(\mathbb{F}_p)$ is a finite abelian group, and we define

$$a_p = p - \#E_p^{\mathrm{ns}}(\mathbb{F}_p).$$

This is completely analogous to the good reduction case, where $a_p = p + 1 - \#E_p(\mathbb{F}_p)$; we have removed the point $(0, 0)$ from consideration, so we should "expect" the cardinality of $E_p^{\mathrm{ns}}(\mathbb{F}_p)$ to be p, rather than $p + 1$, and a_p measures the deviation from this value.

There are two cases to consider, depending on whether $f(x)$ has a double or triple root at 0, and these give rise to three possibilities for the group $E_p^{\mathrm{ns}}(\mathbb{F}_p)$.

- **Case 1: triple root** ($y^2 = x^3$)

  We have the projective curve $zy^2 = x^3$. After removing the singular point $(0 : 0 : 1)$, every other projective point has non-zero y coordinate, so we can normalize the points so that $y = 1$, and work with the affine curve $z = x^3$. There are p-solutions to this equation (including $x = 0$ and $z = 0$, which corresponds to the projective point $(0 : 1 : 0)$ at infinity on our original curve). It follows that $E_p^{\mathrm{ns}}(\mathbb{F}_p)$ is a cyclic group of order p, which is isomorphic to the additive group of $\mathbb{F}_p$; see [10, §2.10] for an explicit isomorphism. In this case we have $a_p = 0$ and say that E has *additive reduction* at p.

- **Case 2: double root** $y^2 = x^3 + ax^2$, $a \neq 0$.

  We have the projective curve $zy^2 = x^3 + ax^2z$, and the point $(0 : 1 : 0)$ at infinity is the only non-singular point on the curve whose x-coordinate is zero. Excluding the point at infinity for the moment, let us divide both sides by x^2, introduce the variable $t = y/x$, and normalize $z = 1$. This yields the affine curve $t^2 = x + a$, and the number of points with $x \neq 0$ is

  $$\begin{aligned}\sum_{x \neq 0} \left(1 + \left(\frac{x+a}{p}\right)\right) &= -\left(1 + \left(\frac{a}{p}\right)\right) + \sum_x \left(1 + \left(\frac{x+a}{p}\right)\right) \\ &= -\left(1 + \left(\frac{a}{p}\right)\right) + \sum_x \left(1 + \left(\frac{x}{p}\right)\right) \\ &= -\left(1 + \left(\frac{a}{p}\right)\right) + p\end{aligned}$$

  where $\left(\frac{a}{p}\right)$ is the Kronecker symbol. If we now add the point at infinity back in we get a total of $p - \left(\frac{a}{p}\right)$ points, thus $a_p = \left(\frac{a}{p}\right)$.

  In this case we say that E has *multiplicative reduction* at p, and further distinguish the cases $a_p = 1$ and $a_p = -1$ as *split* and *non-split* respectively. One can show that in the former case $E_p^{\mathrm{ns}}(\mathbb{F}_p)$ is isomorphic to the multiplicative group $\mathbb{F}_p^\times$, and in the latter case it is isomorphic to the multiplicative subgroup of $\mathbb{F}_{p^2} = \mathbb{F}_p[x]/(x^2 - a)$ made up by the elements of norm 1; see [10, §2.10].

To sum up, there are three possibilities for $a_p = p - \#E_p^{\mathrm{ns}}(\mathbb{F}_p)$:

$$a_p = \begin{cases} 0 & \text{additive reduction,} \\ +1 & \text{split multiplicative reduction,} \\ -1 & \text{non-split multiplicative reduction.} \end{cases}$$

There is one further issue to consider. It could happen that the reduction type of E at a prime p changes when we consider E as an elliptic curve over a finite extension $K/\mathbb{Q}$ (in which case we are then talking about reduction modulo primes $\mathfrak{p}$ of K lying above p). It turns out that this can only happen when E has additive reduction at p. This leads to the following definition.

Definition 25.27. An elliptic curve $E/\mathbb{Q}$ is *semi-stable* if it does not have additive reduction at any prime.

As we shall see in the next lecture, for the purposes of proving Fermat's Last Theorem, we can restrict our attention to semi-stable elliptic curves, and this simplifies matters.

25.8 *L*-series of elliptic curves versus *L*-series of modular forms

Having define the L-series

$$L_E(s) = \prod_p (L_p(p^{-s}))^{-1} = \sum_{n=1}^{\infty} a_n n^{-s}$$

of an elliptic curve $E/\mathbb{Q}$, we now note that the coefficients a_n satisfy the same recurrence relations as those of a weight-2 eigenform. We have $a_1 = 1$, and, as in Corollary 25.12, we have $a_{mn} = a_m a_n$ for all $m \perp n$, and $a_{p^{r+1}} = a_p a_{p^r} - p a_{p^{r-1}}$ for all primes p of good reduction, as you proved on Problem Set 7. Moreover, for the bad primes we have $a_p \in \{0, \pm 1\}$ and it easy to check that $a_{p^r} = a_p^r$, which also applies to the coefficients of an eigenform in $S_k^{\text{new}}(\Gamma_0(N))$ when $p|N$ (see Remark 25.15).

So now we might ask, given an elliptic curve $E/\mathbb{Q}$, is there a modular form f for which $L_E(s) = L_f(s)$? Or, to put it more simply, let $L_E(s) = \sum_{n=1}^{\infty} a_n n^{-s}$, and define

$$f_E(\tau) = \sum_{n=1}^{\infty} a_n q^n.$$

Our question then becomes: is $f_E(\tau)$ a modular form?

It's clear from the recurrence relation for a_{p^r} that if $f_E(\tau)$ is a modular form, then it must be a modular form of weight 2; but there are additional constraints. For $k = 2$ the equations (4) and (5) both give the Euler product

$$\prod_p \left(1 - a_p p^{-s} + \chi(p) p p^{-2s}\right)^{-1},$$

and it is essential that $\chi(p)$ is the same in both cases; recall that for eigenforms f in $S_k^{\text{new}}(\Gamma_0(N))$ we have $\chi(p) = 0$ for primes $p|N$, while for elliptic curves $E/\mathbb{Q}$ we have $\chi(p) = 0$ for primes $p|\Delta_{\min}(E)$. No elliptic curve over $\mathbb{Q}$ has good reduction at every prime, so we cannot use eigenforms of level 1, we need to consider newforms of some level N.

This suggests taking N to be the product of the prime divisors of $\Delta_{\min}(E)$, but we should note that any N with the same set of prime divisors would have the same property. In turns out that for semi-stable elliptic curves, simply taking the product of the prime divisors of $\Delta_{\min}(E)$ is the right thing to do, and this is all we need for the proof of Fermat's Last Theorem.

Definition 25.28. Let $E/\mathbb{Q}$ be a semi-stable elliptic curve with minimal discriminant $\Delta_{\min}$. The *conductor* N_E of E is the product of the prime divisors of $\Delta_{\min}$.

Remark 25.29. For elliptic curves that are not semistable, at primes $p > 3$ where E has additive reduction we simply replace the factor p in N_E by p^2. But the primes 2 and 3 require special treatment (as usual), and the details can get quite technical; see [9, IV.10]. In any case, the conductor of an elliptic curve $E/\mathbb{Q}$ is squarefree if and only if it is semistable.

We can now say precisely what it means for an elliptic curve over $\mathbb{Q}$ to be modular.

Definition 25.30. $E/\mathbb{Q}$ is *modular* if $f_E(\tau)$ is a modular form of weight 2 for $\Gamma_0(N_E)$.

In the case that E/Q is modular, the modular form f_E will necessarily be a newform. As noted at the beginning of the lecture, we now know that this is always the case.

Theorem 25.31 (Modularity Theorem, formerly the Shimura-Taniyama-Weil[8] conjecture)**.** *Every elliptic curve $E/\mathbb{Q}$ is modular.*

25.9 BSD and the parity conjecture

When E is modular, the L-series of E and the modular form f_E necessarily coincide, and this implies that $L_E(s)$ has an analytic continuation and satisfies a functional equation, since this holds for the L-series of a modular form, by Theorem 25.22. But prior to the modularity theorem, this was an open question known as the Hasse-Weil conjecture.

Theorem 25.32. *Let E be an elliptic curve over $\mathbb{Q}$. Then $L_E(s)$ has an analytic continuation to a meromorphic function on $\mathbb{C}$, and*

$$\tilde{L}_E(s) = N_E^{s/2}(2\pi)^{-s}\Gamma(s)L_E(s)$$

satisfies the functional equation

$$\tilde{L}_E(s) = w_e\tilde{L}_E(2-s),$$

where $w_e = \pm 1$.

The sign w_E in the functional equation is called the *root number* of E. If $w_E = -1$ then the functional equation implies that $\tilde{L}_E(s)$, and therefore $L_E(s)$, has a zero at $s = 1$; in fact it is not hard to show that $w_E = 1$ if and only if $L_E(s)$ has a zero of even order at $s = 1$.

The conjecture of Birch and Swinnerton-Dyer (BSD) relates the behavior of $L_E(s)$ at $s = 1$ to the rank of $E(\mathbb{Q})$. Recall that

$$E(\mathbb{Q}) \simeq E(\mathbb{Q})_{\text{tor}} \times \mathbb{Z}^r,$$

where $E(\mathbb{Q})_{\text{tor}}$ denotes the torsion subgroup of $E(\mathbb{Q})$ and r is the *rank* of E.

Conjecture 25.33 (Weak BSD conjecture)**.** *Let $E/\mathbb{Q}$ be an elliptic curve of rank r. Then $L_E(s)$ has a zero of order r at $s = 1$.*

The strong version of the BSD conjecture makes a more precise statement, but a proof of even the weak version is enough to claim the Millennium Prize for the BSD conjecture. There is also the weaker parity conjecture, which relates the root number w_E in the functional equation to the parity of r.

Conjecture 25.34 (Parity conjecture)**.** *Let $E/\mathbb{Q}$ be an elliptic curve of rank r. Then the root number is given by $w_E = (-1)^r$.*

[8] One can find references to this conjecture in the literature that use just two of these three names (which two depends on the author; all three possibilities occur and some even order the names differently). Thankfully, the conjecture has been proved and everyone is now happy to call it the Modularity Theorem.

25.10 Modular elliptic curves

The relationship between elliptic curves and modular forms is remarkable and not at all obvious. It is reasonable to ask why people believed the modular conjecture in the first place. The most compelling reason is that every newform of weight 2 gives rise to a modular elliptic curve.

Theorem 25.35 (Eichler-Shimura)**.** *Let $f = \sum a_n q^n$ be a weight 2 newform for $\Gamma_0(N)$ with $a_n \in \mathbb{Z}$. Then there exists an elliptic curve $E/\mathbb{Q}$ of conductor N for which $f_E = f$.*

See [6, V.6] for details on how to construct the elliptic curve given by the theorem, which was known long before the modularity theorem was proved.

The elliptic curve E whose existence is guaranteed by the Eichler-Shimura theorem is only determined up to isogeny.[9] This is due to the fact that isogenous elliptic curves E and E' over $\mathbb{Q}$ necessarily have the same L-series, and therefore $f_E = f_{E'}$. It is easy to show that the a_p values in the L-series of E and E' must agree at every prime p at which both curves have good reduction (all but finitely many primes), since the trace of Frobenius is preserved by isogenies, and it turns out that in fact E and E' must have the same reduction type at every prime so their L-series are actually identical. The converse also holds, but this is not so easy to show. In fact, something even stronger is true [6, Thm. V.4.1].

Theorem 25.36 (Tate-Faltings)**.** *Let E and E' be elliptic curves over $\mathbb{Q}$ with L-series $L_E(s) = \sum a_n n^{-s}$ and $L_{E'}(s) = \sum a'_n n^{-s}$, respectively. If $a_p = a'_p$ for sufficiently many primes p of good reduction for E and E', then E and E' are isogenous.*

What "sufficiently many" means depends on the curves E and E', but the key point is that it is a finite number. For any positive integer N, one can enumerate all the newforms in $S_2^{new}(\Gamma_0(N))$ with integral q-expansions; this is a finite list. It is also possible (but not easy) to enumerate all the isogeny classes of elliptic curves with conductor N; this is also a finite list. When this was done for various small values of N, it was found that the two lists matched perfectly in every case. It was this matching that made the modularity conjecture truly compelling.

References

[1] A. Agashe, K. Ribet, and W.A. Stein, *The Manin constant*, Pure and Applied Mathematics Quarterly **2** (2006), 617–636.

[2] A.O.L. Atkin and L. Lehner, *Hecke operators on $\Gamma_0(m)$*, Mathematische Annalen **185** (1970), 134–160.

[3] C. Breuil, B. Conrad, F. Diamond, and R. Taylor, *On the modularity of elliptic curves over $\mathbb{Q}$: wild 3-adic exercises*, Journal of the American Mathematical Society **14** (2001), 843–939.

[4] F. Diamond and J. Shurman, *A first course in modular forms*, Springer, 2005.

[5] M. Laska, *An algorithm for finding a minimal Weierstrass equation for an elliptic curve*, Mathematics of Computation **38** (1982), 257-260.

[9] There is an "optimal" representative for each isogeny class; see John Cremona's appendix to [1].

[6] J.S. Milne, *Elliptic curves*, BookSurge Publishers, 2006.

[7] J.-P. Serre, *A course in arithmetic*, Springer, 1973.

[8] J.H. Silverman, *The arithmetic of elliptic curves*, second edition, Springer, 2009.

[9] J.H. Silverman, *Advanced topics in the arithmetic of elliptic curves*, Springer, 1994.

[10] L.C. Washington, *Elliptic Curves: Number Theory and Cryptography*, second edition, Chapman and Hall/CRC, 2008.

[11] R. Taylor and A. Wiles, *Ring-theoretic properties of certain Hecke algebras*, Annals of Mathematics **141** (1995), 553–572.

[12] A. Wiles, *Modular elliptic curves and Fermat's last theorem*, Annals of Mathematics **141** (1995), 443-551.