18.783 Elliptic Curves
Lecture 24

Andrew Sutherland

May 4, 2022
The modularity theorem

Definition
An elliptic curve $E/\mathbb{Q}$ is modular if it has the same $L$-function as a modular form.

Theorem (Taylor-Wiles 1995)
Every semistable elliptic curve $E/\mathbb{Q}$ is modular.

Corollary (Wiles 1995)
The equation $x^n + y^n = z^n$ has no integers solutions with $xyz \neq 0$ for $n > 2$.

Theorem (Breuil-Conrad-Diamond-Taylor 2001)
Every elliptic curve $E/\mathbb{Q}$ is modular.
Weak modular forms

**Definition**

A holomorphic function \( f : \mathcal{H} \to \mathbb{C} \) is a **weak modular form** of weight \( k \) for a congruence subgroup \( \Gamma \) if for every \( \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma \) we have

\[
f(\gamma \tau) = (c \tau + d)^k f(\tau).
\]

If \(-I \in \Gamma\), for odd \( k \) the only weak modular form of weight \( k \) is the zero function.

**Example**

The \( j \)-function \( j(\tau) \) is a weak modular form of weight 0 for \( \text{SL}_2(\mathbb{Z}) \), and for \( k \geq 3 \)

\[
G_k(\tau) := G_k([1, \tau]) := \sum_{\substack{m,n \in \mathbb{Z} \\ (m,n) \neq (0,0)}} \frac{1}{(m + n\tau)^k},
\]

is a weak modular form of weight \( k \) for \( \text{SL}_2(\mathbb{Z}) \).
Modular forms

If $\Gamma(N) \subseteq \Gamma$ then $f(\tau + N) = f(\tau)$ for any weak modular form $f : \mathcal{H} \to \mathbb{C}$. It follows that $f$ has a $q$-expansion (at $\infty$) of the form

$$f(\tau) = f^*(q^{1/N}) \sum_{n=-\infty}^{\infty} a_n q^{n/N} \quad (q := e^{2\pi i \tau})$$

Definition

A weak modular form $f$ is holomorphic at $\infty$ if $f^*$ is holomorphic at 0, and $f$ is holomorphic at the cusps if $f(\gamma \tau)$ is holomorphic at $\infty$ for all $\gamma \in \text{SL}_2(\mathbb{Z})$. A modular form is a weak modular form that is holomorphic at the cusps.

Example

The $j$-function is not a modular form, but the Eisenstein series $G_k(\tau)$ is a modular form of weight $k$ for all even $k \geq 4$. 
Cusp forms

**Definition**

A modular form is a **cusp form** if it vanishes at all the cusps; equivalently its \( q \)-expansion has the form \( \sum_{n \geq 1} a_n q^n \) (at every cusp).

**Example**

The Eisenstein series \( G_k(\tau) \) are not cusp forms but the discriminant function

\[
\Delta(\tau) = g_2(\tau)^3 - 27g_3(\tau)^2
\]

is a cusp form of weight 12 for \( \text{SL}_2(\mathbb{Z}) \).

The set \( M_k(\Gamma) \) of modular forms of weight \( k \) for \( \Gamma \) is a \( \mathbb{C} \)-vector space that contains the set of cusp forms \( S_k(\Gamma) \) as a subspace. For \( k = 2 \) we have \( \dim S_k(\Gamma) = g(\Gamma) \).
Hecke operators

Definition

For $n \in \mathbb{Z}_{>0}$ the Hecke operator (or Hecke correspondence) $T_n$ is a linear operator on the free abelian group of lattices $L := \langle \omega_1, \omega_2 \rangle$ defined by

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

We also define the homethety operator $R_\lambda$ by $L \mapsto \lambda L$, for all $\lambda \in \mathbb{C}^\times$.

Theorem

The operators $T_n$ and $R_\lambda$ 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) $T_{pr+1} = T_p T_p - p T_{pr} R_p$ for all primes $p$ and integers $r \geq 1$. 
The action of Hecke operators on modular forms

Each modular form $f: \mathcal{H} \to \mathbb{C}$ of weight $k$ defines a function on lattices $[\omega_1, \omega_2]$ via

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

**Definition**

For $f \in M_k(\Gamma_0(1))$ we define

$$R_\lambda f(\tau) := f(\lambda[1, \tau]) = \lambda^{-k} f(\tau) \in M_k(\Gamma_0(1)),$$

$$T_n f(\tau) := n^{k-1} \sum_{[[1, \tau]: L] = n} f(L) = n^{k-1} \sum_{ad = n, 0 \leq b < d} d^{-k} f\left(\frac{a\tau + b}{d}\right) \in M_k(\Gamma_0(1)).$$

$R_\lambda$ and $T_n$ are linear operators on $M_k(\Gamma_0(1))$ that we can restrict to $S_k(\Gamma_0(1))$. We have $T_{mn} = T_m T_n$ for $m \perp n$, and $T_{p^{r+1}} = T_p T_p - p^{k-1} T_{p^{r-1}}$ for $p$ prime.
**Theorem**

For any \( f \in S_k(\Gamma_0(1)) \) and prime \( p \) we have

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

and for all \( m \perp n \) we have \( a_m(T_n f) = a_{mn}(f) \). In particular \( a_1(T_n f) = a_n(f) \).

**Definition**

An eigenform for \( S_k(\Gamma_0(1)) \) satisfies \( T_n f = \lambda_n f \) for some \( \lambda_1, \lambda_2, \ldots \in \mathbb{C}^\times \).

We normalize eigenforms so that \( a_1(f) = 1 \), and then \( \lambda_n = a_n \) for all \( n \in \mathbb{Z}_{>0} \).

We then have \( a_m a_n = a_{mn} \) for \( m \perp n \) and \( a_p^r = a_p a_p^{r-1} - p^{k-1}a_p^{r-2} \) for \( p \) prime.
A basis of eigenforms

Definition

Let $\Gamma$ be a congruence subgroup. The Petersson inner product on $S_k(\Gamma)$ is defined by

$$\langle f, g \rangle = \int_F f(\tau)\overline{g(\tau)}y^{k-2}dxdy.$$

It is a positive definite Hermitian form on $S_k(\Gamma)$: it is bilinear and $\langle f, g \rangle = \langle g, f \rangle$, with $\langle f, f \rangle = 0$ if and only if $f = 0$. Moreover, we have $\langle f, T_ng \rangle = \langle T_nf, g \rangle$. The Hecke operators are thus Hermitian operators on the space $S_k(\Gamma)$.

Theorem

The space of cusp forms for $S_k(\Gamma_0(1))$ is a direct sum of one-dimensional Hecke eigenspaces, and it has a unique basis of normalized eigenforms $f(\tau) = \sum a_nq^n$ for which $a_n$ is the eigenvalue of $T_n$ on the subspace spanned by $f$. 
The Atkin-Lehner theory of newforms

Definition
A cusp form $f \in S_k(\Gamma_0(N))$ is old if $f \in S_k(\Gamma_0(M))$ for some $M$ properly dividing $N$. The set of old forms is a subspace $S_k^{\text{old}}(\Gamma_0(N))$ of $S_k(\Gamma_0(N))$. Taking the orthogonal complement with respect to the Petersson inner product yields

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

The level of $f \in S_k(\Gamma_0(N))$ is the unique $M|N$ for which $f \in S_k^{\text{new}}(\Gamma_0(M))$. Normalized eigenforms $f \in S_k^{\text{new}}(\Gamma_0(N))$ are called newforms.

Theorem (Atkin-Lehner)

The space $S_k^{\text{new}}(\Gamma_0(N))$ is a direct sum of one-dimensional Hecke eigenspaces, each generated by a newform $f(\tau) = \sum_n a_n q^n$ for which $a_n$ is the eigenvalue of $T_n$ on $\langle f \rangle$. 
Dirichlet series

**Definition**

A **Dirichlet series** is a function of the form \( L(s) = \sum_{n \geq 1} a_n n^{-s} \) with \( a_n \in \mathbb{C} \).

If \( |a_n| = O(n^\sigma) \) then \( L(s) \) converges locally uniformly in the half plane \( \text{re}(s) > 1 + \sigma \).

**Example**

The **Riemann zeta function** is the Dirichlet series \( \zeta(s) = \sum_{n \geq 1} n^{-s} \).

It converges locally uniformly to a holomorphic function on \( \text{re}(s) > 1 \), with a simple pole at \( s = 1 \) and no other poles. Moreover, the following hold:

- \( \zeta(s) \) has an **analytic continuation** to a meromorphic function on \( \mathbb{C} \);
- \( \tilde{\zeta}(s) = \pi^{-s/2} \Gamma \left( \frac{s}{2} \right) \zeta(s) \) satisfies\(^1\) the **functional equation** \( \hat{\zeta}(s) = \hat{\zeta}(1 - s) \);
- we have the **Euler product** \( \zeta(s) = \prod_p (1 - p^{-s})^{-1} \).

\[^1\text{Here } \Gamma(s) := \int_0^\infty e^{-t} t^{s-1} \, dt \text{ is the Euler gamma function.}\]
**$L$-functions of modular forms**

**Definition**

The *L-function* of a cusp form $f = \sum a_n q^n$ is the Dirichlet series $L(f, s) := \sum a_n n^{-s}$. If $f$ has weight $k$ then $L(f, s)$ converges locally uniformly on $\text{re}(s) > 1 + k/2$.

**Theorem (Hecke)**

For $f \in S_k(\Gamma_0(N))$ the $L$-function $L(f, s)$ has an holomorphic continuation to $\mathbb{C}$ and \( \hat{L}(f, s) := N^{s/2}(2\pi)^{-s}\Gamma(s)L(f, s) \) satisfies \( \hat{L}(f, s) = \pm \hat{L}(f, k-s) \).

For $f \in S_k^{\text{new}}(\Gamma_0(N))$ the $L$-function $L(f, s)$ has the Euler product

\[
L(f, s) = \sum_{n \geq 1} a_n n^{-s} = \prod_p (1 - a_p p^{-s} + \chi(p) p^{k-1} p^{-2s})^{-1},
\]

where the Dirichlet character $\chi$ satisfies $\chi(p) = 0$ for $p|N$ and $\chi(p) = 1$ otherwise.
Summary of modular forms for $\Gamma_0(N)$

- A modular form of weight $k$ for $\Gamma_0(N)$ is a holomorphic function $f : \mathcal{H}^* \to \mathbb{C}$ satisfying $f(\gamma \tau) = (c\tau + d)^k f(\tau)$ for all $\gamma = \left( \begin{array}{cc} a & b \\ c & d \end{array} \right) \in \Gamma_0(N)$.

- A cusp form $f \in S_k(\Gamma_0(N))$ vanishes at the cusps (its $q$-expansion has $a_0 = 0$).

- The cusp forms $S_k(\Gamma_0(N))$ are a $\mathbb{C}$-vector space with a Petersson inner product.

- The Hecke operators $T_n$ are commuting Hermitian operators on $S_k(\Gamma_0(N))$.

- An eigenform $f = \sum a_n q^n \in S_k(\Gamma_0(N))$ satisfies $T_n f = a_n f$ for all $n \geq 1$.

- A cusp form $f \in S_k \Gamma_0(N)$ is old if $f \in S_k(\Gamma_0(M))$ for some proper divisor $M | N$, and we have $S_k(\Gamma_0(N)) = S^\text{old}_k(\Gamma_0(N)) \oplus S^\text{new}_k(\Gamma_0(N))$.

- The level of $f \in S_k(\Gamma_0(N))$ is the least $M | N$ for which $f \in S^\text{new}_k(\Gamma_0(M))$.

- The newforms of weight $k$ and level $N$ are a canonical basis for $S^\text{new}_k(\Gamma_0(N))$.

- The $L$-function $L(f, s)$ has an analytic continuation, a functional equation satisfied by $\hat{L}(f, s)$, and an Euler product $\prod (1 - a_p p^{-s} + \chi(p)p^{k-1}p^{-2s})^{-1}$. 

The $L$-function of an elliptic curve over $\mathbb{Q}$

**Definition**

The $L$-function of an elliptic curve $E/\mathbb{Q}$ is defined by the Euler product

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

where $\chi(p)$ is 0 if $E$ has **bad reduction** at $p$, and 1 otherwise. For primes of good reduction $a_p := p + 1 - \#E(\mathbb{F}_p)$ is the trace of Frobenius, and otherwise

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

This means that $a_p \in \{0, \pm1\}$ at bad primes.
Primes of bad reduction

Definition

Let $K$ be a number field. An integral model for $E/K$ is a Weierstrass equation

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

with $a_1, a_2, a_3, a_4, a_6 \in \mathcal{O}_K$. The minimal discriminant of $E/K$ is the $\mathcal{O}_K$-ideal

$$\Delta_{\min}(E) := \prod_p p^{\min v_p(\Delta)}$$

where $p$ varies over primes of $K$ and $\Delta$ over discriminants of integral models for $E$. A prime of bad reduction for $E$ is a prime $p$ of $K$ that divides the ideal $\Delta_{\min}(E)$.

A global minimal model for $E/K$ is an integral model with discriminant $\Delta_{\min}(E)$. Such models always exist when $K$ has class number one (and in particular for $K = \mathbb{Q}$).
Why we like (general) Weierstrass equations

Every elliptic curve $E/\mathbb{Q}$ can be defined by an equation of the form $y^2 = x^3 + Ax + B$.

But equations of this form are usually not global minimal models, and a prime $p$ that divides the discriminant $-16(4A^3 + 27B^2)$ is not necessarily a prime of bad reduction, even though $y^2 = x^3 + Ax + B$ defines a singular curve over $\mathbb{F}_p$ in this case.

Example

Consider the elliptic curve $y^2 = x^3 - 13392x - 1080432$ over $\mathbb{Q}$.
We have $A = 2^4 \cdot 3^3 \cdot 31$ and $B = 2^4 \cdot 3^3 \cdot 41 \cdot 61$ (so no extraneous powers), and

$$\Delta = -16(4A^3 + 27B^2) = -350572971995136 = -2^{12}3^{12}11^5.$$ 

But 2 and 3 are not primes of bad reduction!
Indeed, $y^2 + y = x^3 - x^2$ is a global minimal model with discriminant $\Delta_{\text{min}}(E) = -11$. 
Types of bad reduction

If \( p \) is an odd prime of bad reduction for \( E / \mathbb{Q} \) we can find an integral model \( y^2 = f(x) \) whose discriminant \( \Delta \) satisfies \( v_p(\Delta) = v_p(\Delta_{\text{min}}) > 0 \), and \( f(x) \) then has a repeated root \( r \) modulo \( p \). Without loss of generality, we assume \( r = 0 \) (replace \( x \) with \( x - r \)).

Over \( \mathbb{F}_p \) we then have the curve \( \overline{E} : y^2z = x^3 + ax^2z \) with a singular point \((0 : 0 : 1)\). Now define \( E^{\text{ns}}(\mathbb{F}_p) := E(\mathbb{F}_p) - \{(0 : 0 : 1)\} \) and let \( a_p := p - \#E^{\text{ns}}(\mathbb{F}_p) \in \mathbb{Z} \).

The set \( E^{\text{ns}}(\mathbb{F}_p) \) is a finite abelian group (under the usual group law) and we have

<table>
<thead>
<tr>
<th>( \frac{a}{p} )</th>
<th>( #E^{\text{ns}}(\mathbb{F}_p) )</th>
<th>( E^{\text{ns}}(\mathbb{F}_p) )</th>
<th>reduction type</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>( p )</td>
<td>( \simeq \mathbb{F}_p )</td>
<td>additive</td>
</tr>
<tr>
<td>+1</td>
<td>( p - 1 )</td>
<td>( \simeq \mathbb{F}_p^\times )</td>
<td>split multiplicative</td>
</tr>
<tr>
<td>−1</td>
<td>( p + 1 )</td>
<td>( \simeq { \alpha \in \mathbb{F}_p^\times : \alpha^{p+1} = 1 } )</td>
<td>non-split multiplicative</td>
</tr>
</tbody>
</table>

Note that \( a_p = p - \#E^{\text{ns}}(\mathbb{F}_p) = \left( \frac{a}{p} \right) \) in every case. Something similar works for \( p = 2 \).
The conductor of an elliptic curve

Definition

The conductor of an elliptic curve $E/\mathbb{Q}$ is the integer $N_E := \prod_p p^{\varepsilon(p) + \delta(p)}$

where $\varepsilon(p) = 0, 1, 2$ when $E$ has good, multiplicative, additive reduction at $p$.

The “wild” exponent $\delta(p)$ is zero unless we have additive reduction at $p = 2, 3$ in which case it can be defined using the ramification of $p$ in the $p^n$-torsion fields $\mathbb{Q}(E[p^n])$.

We have $N_E | \Delta_{\min}(E)$ with $v_p(N_E) \leq 8, 5$ for $p = 2, 3$ and $v_p(N_E) \leq 2$ for $p > 3$.

Definition

An elliptic curve $E/\mathbb{Q}$ is semistable if its conductor is squarefree.

Equivalently, $E$ does not have additive reduction at any prime.
Modularity

**Definition**

For an elliptic curve $E/\mathbb{Q}$ with $L(E, s) = \sum a_n n^{-s}$ we define $f_E : \mathcal{H} \rightarrow \mathbb{C}$ by

$$f_E(\tau) := \sum_{n \geq 1} a_n q^n \quad (q := e^{2\pi i \tau})$$

The elliptic curve $E$ is **modular** if the function $f_E$ is a modular form. Equivalently, $E$ is modular if and only if $L(E, s)$ is the $L$-function of a modular form.

If $E$ is modular then $f_E$ must be a cusp form of weight 2 since the Euler factors are

$$1 - a_p p^{-s} + \chi(p) p p^{-2s} = 1 - a_p p^{-s} + \chi(p) p^{k-1} p^{-2s},$$

**Theorem (Modularity theorem)**

Let $E/\mathbb{Q}$ be an elliptic curve. Then $f_E$ is an eigenform of weight 2 and level $N_E$. 
The functional equation

**Corollary**

Let $E/\mathbb{Q}$ be an elliptic curve. The $L$-function $L(E, s)$ has a holomorphic continuation to $\mathbb{C}$ and $\hat{L}(E, s) := N_E^{s/2}(2\pi)^{-s}\Gamma(s)L_E(s)$ satisfies $\hat{L}(E, s) = \pm \hat{L}(E, 2 - s)$.

Notice that $\hat{L}(E, s) = -\hat{L}(E, 2 - s)$ is possible only when $\text{ord}_{s=1}L(E, s)$ is odd.

**Conjecture (Weak BSD)**

We have $E(\mathbb{Q}) \simeq \mathbb{Z}^r \oplus E(\mathbb{Q})_{\text{tors}}$ if and only if $\text{ord}_{s=1}L(E, s) = r$.

**Conjecture (Parity conjecture)**

If $E(\mathbb{Q}) \simeq \mathbb{Z}^r \oplus E(\mathbb{Q})_{\text{tors}}$ then $\hat{L}(E, s) = (-1)^r\hat{L}(E, 2 - s)$. 
Eichler-Shimura

**Definition**

Let $f = \sum a_n q^n \in S_{2}^{\text{new}}(\Gamma_0(N))$ be a newform. The coefficients $a_n$ are algebraic integers that generate a finite extension $\mathbb{Q}(f)/\mathbb{Q}$. The dimension of $f$ is $\dim f := [\mathbb{Q}(f) : \mathbb{Q}]$; we call $f$ rational if $\dim f = 1$.

One can associate to any newform in $f \in S_{2}^{\text{new}}(\Gamma_0(N))$ a lattice $\Lambda$ in $\mathbb{C}^d$ and a corresponding abelian variety $A_f := \mathbb{C}^d/\Lambda$ of dimension $d = \dim f$ defined over $\mathbb{Q}$. One then has $L(A, s) = \prod_\sigma L(\sigma(f), s)$ where $\sigma(f)$ ranges over the $\text{Aut}(\mathbb{C})$-orbit of $f$ (equivalently, $a_n \in \mathbb{Q}(f)$ and $\sigma$ varies over embeddings of $\mathbb{Q}(f)$ into $\mathbb{C}$).

**Theorem (Eichler-Shimura, Carayol)**

*For every rational newform $f \in S_{2}^{\text{new}}(\Gamma_0(N))$ there is an elliptic curve $E/\mathbb{Q}$ of conductor $N$ with $f_E = f$ and $L(E, s) = L(f, s)$.***
Faltings-Tate

Recall that isogenous elliptic curves over $\mathbb{F}_p$ have the same trace of Frobenius. If $E_1$ and $E_2$ are isogenous elliptic curves over $\mathbb{Q}$, then $a_p(E_1) = a_p(E_2)$ for all primes of good reduction, and in fact $a_p(E_1) = a_p(E_2)$ for all primes.

It follows that isogenous elliptic curves over $\mathbb{Q}$ have the same $L$-function. Remarkably, the converse holds, in fact something even stronger holds.

Theorem (Faltings-Tate)

If two elliptic curves $E, E'$ over $\mathbb{Q}$ satisfy $a_p(E) = a_p(E')$ for all but finitely many primes $p$ then $E$ and $E'$ are isogenous (thus $a_p(E) = a_p(E')$ for all primes $p$).

Corollary

Elliptic curves over $\mathbb{Q}$ are isogenous if and only if they have the same $L$-function.
Isogeny classes of elliptic curves and modular forms

Distinct eigenforms $S^\text{new}_2(\Gamma_0(N))$ necessarily have distinct $L$-functions, since their $q$-expansions $\sum a_n q^n$ must be linearly independent. The modular form $f_E$ given by the modularity theorem thus depends only on the isogeny class of $E/\mathbb{Q}$ and in general there may be non-isomorphic isogenous $E/\mathbb{Q}$ that correspond to the same $f_E$.

There is thus in general a many-to-one relationship between elliptic curves over $\mathbb{Q}$ and rational eigenforms of weight 2, but a one-to-one relationship between isogeny classes of elliptic curves over $\mathbb{Q}$ and rational eigenforms of weight 2.

You can see this explicitly in the $L$-functions and Modular Forms Database (LMFDB).

Example

The elliptic curves 11.a1, 11.a2, 11.a3 of conductor $N_E = 11$ make up the isogeny class 11.a, which corresponds to the modular form 11.2.a.a of weight 2 and level 11. They all have the same $L$-function 2-11-1.1-c1-0-0, which has $\text{ord}_{s=1} L(s) = 0$. 