18.784: Seminar in Number Theory

Lecturer: Professor Ju-Lee Kim

Notes by: Andrew Lin

Spring 2020

Introduction

This class is a CI-M class, so it will consist primarily of student presentations.

Looking at the syllabus, most of this class is centered around presentation, writing, and participation. The schedule is already posted on the Stellar website, and there is a Google Sheet where we can sign up for presentation times. Our first talk is just for feedback – it won't be graded.

We'll be starting with chapter 7 of Serre's "A Course in Arithmetic," and then we'll move on to Diamond and Shurman's "A First Course in Modular Forms." We can get both of these texts from the library.

Fact 1

Notes from my own presentations will be copied from my handouts, while notes from others' are my transcriptions.

1 February 4, 2020

Here's a few points about how to give a good presentation:

- Read and understand (digest) the material we're being given, and then explain it! The rest of the class will only see this final "explain" part, but that doesn't mean the first two parts aren't important.
- Have a good lecture plan. Decide what to include and what to exclude, keeping in mind that we're trying to help our classmates understand the material. Also, include good examples!
- Be ready to respond to questions. We'll have lots of discussions, and we get to ask many questions.
- Speak clearly and loudly.
- Organize board space well (dividing the board into smaller boards, picking blackboard order, etc.) Make sure the lecture "flows" well on the board, so people can go back to previous theorem statements, definitions, and so on. (Write in complete sentences.)
- Label definitions, lemmas, and so on.
- Make handouts for lectures and distribute them beforehand. (This also helps with bad handwriting.) Handouts may or may not be graded, still unclear.
- Interact with the audience – face them from time to time when lecturing, make eye contact, and pause for questions.

Professor Kim has office hours (see Stellar for timeslots each week), where she can help us with presentation plans and material. All of us will be given comment forms to fill out for each presenter – we should make sure our comments are constructive, because they will be given (anonymously) to the presenter.

For our writing project, we can pick our own topic – Professor Kim is very open-minded.

Our main goals in this class are to understand modular forms, elliptic curves, and L-functions. So today, we'll do a brief introduction for motivation and to tie together some of these topics.

Example 2

The **Riemann zeta function**

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

is an example of an **L-function**.

This function is well-defined for all $\text{Re}(s) > 1$, and it has a pole at $s = 1$. We can also write it in the Euler product form

$$\zeta(s) = \prod_p \frac{1}{1 - p^{-s}}.$$

$\zeta(s)$ also satisfies a functional equation, which gives a **meromorphic continuation** to all of $\mathbb{C}$. The whole idea is that this can contain a lot of information!

Definition 3 (Loose definition)

In general, an **L-function** takes the form

$$L(s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s},$$

where $a_n \in \mathbb{C}$ and the sequence $\{a_n\}$ contains some arithmetic information.

Often, L-functions can be associated with an elliptic curve $L(s, E)$ or a modular form $L(s, f)$ – we'll talk about this more later.

Example 4 (Dirichlet series)

Let $N \in \mathbb{Z}_{>0}$, and consider a **character** (basically a one-dimensional representation) from $(\mathbb{Z}/n\mathbb{Z})^* \to \mathbb{C}$.

This can be lifted to $\tilde{\chi} : \mathbb{Z} \to \mathbb{C}$ by defining

$$\tilde{\chi}(a) = \begin{cases} \chi(a \bmod N & \gcd(a, N) = 1 \\ 0 & \text{otherwise} \end{cases}.$$

We can then define the L-function

$$L(s, \chi) = \sum_{n=1}^{\infty} \frac{\tilde{\chi}(n)}{n^s}.$$

This actually gives us the following result from chapter 6 of Serre's book:

Theorem 5 (Dirichlet)

If $\gcd(a, N) = 1$, then there exist infinitely many primes p with $p \equiv a \bmod N$.

Definition 6

An **elliptic curve** is a smooth projective algebraic curve of genus 1 with a distinguished point.

Rational elliptic curves can be written in the form (after transformations)

$$y^2 = x^3 + ax + b, a, b \in \mathbb{Z}$$

where the discriminant $4a^3 + 27b^2$ is nonzero. This is an example of a **Diophantine equation**, which people have been interested in trying to solve for integer solutions:

Example 7

Are there three consecutive integers whose product is a square?

(This can be rewritten as finding integral solutions to $y^2 = x^3 - x$.)

Example 8

Similarly, what positive integers are **congruent** (which means they're a possible area of a right triangle with rational side lengths)?

This can actually be rephrased in terms of elliptic curves as well, after many changes of variables:

Proposition 9

An integer $n \in \mathbb{N}$ is congruent if and only if $y^2 = x^3 - n^2x$ has rational solutions with $y \neq 0$.

For example, for $n = 1$, we have the elliptic curve $y^2 = x^3 - x$, which only has the solutions $(0, 0), (1, 0), (-1, 0)$. So 1 is not a congruent number.

We can be more precise with this problem, though. Denote C_n to be the set of triples $(a, b, c) \in \mathbb{Q}^3$ corresponding to right triangles with area n, and let E_n be the solutions $(x, y) \in \mathbb{Q}^2$ to the elliptic curve $y^2 = x^3 - n^2x$. Then there's a direct bijection between the solutions:

$$(a, b, c) \iff \left(\frac{nb}{c-a}, \frac{2n^2}{c-a}\right), \quad \left(\frac{x^2-n^2}{y}, \frac{2nx}{y}, \frac{x^2+n^2}{y}\right) \iff (x, y).$$

Elliptic curves can also be used to prove Fermat's last theorem – specifically, the relevant objects are called **Frey curves**. For any odd prime p, the idea is to consider

$$y^2 = x(x - a^p)(x + b^p)$$

if $a^p + b^p = c^p$. (Then the discriminant is divisible by a^p, b^p, c^p.) Frey curves are not modular, but it turns out all rational elliptic curves are modular! So this is a contradiction, but it takes many pages to prove this.

So how do we construct L-functions associated to an elliptic curve E? We start with our basic form $\sum_{n=1}^{\infty} \frac{a_n}{n^s}$, and we specify our coordinates: a_p for prime p is equal to $p + 1$ minus the number of rational points of E over $\mathbb{F}_p$, and similarly a_{p^n} is $p^n + 1$ minus the number of rational points of E over $\mathbb{F}_{p^n}$. And then we can define $a_{mn} = a_m \cdot a_n$ for relatively prime m, n, and this allows us to determine all coefficients.

Remark 10. *If we're interested in statistics or probability, we can consider the distribution of the coefficients* $\frac{a_p}{2\sqrt{p}}$ *(which are contained in [-1, 1]). This is called the* ***Sato-Tate distribution****.*

Definition 11

The **upper-half plane** is defined as $\mathbb{H} = \{z \in \mathbb{C} | \text{Im}(z) > 0\}$, and the special linear group $SL_2(\mathbb{R})$ is defined as the set of matrices $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ with real coefficients and determinant 1.

$SL_2(\mathbb{R})$ acts on the upper half-plane by taking a complex number z and sending it to $gz = \frac{az+b}{cz+d}$.

Definition 12

A function $f : \mathbb{H} \to \mathbb{C}$ is a **modular function** of weight $k \in 2\mathbb{N}$ if

$$f(gz) = (cz + d)^k f(z).$$

If f is also holomorphic on $\mathbb{H}$ and at infinity, then f is a **modular form**.

Let's think about $\mathbb{H}/SL_2(\mathbb{Z})$: this is generated by the two elements $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, and this gives us a **fundamental domain**. And if we take the closure, this gives us a **modular curve**, which is closely related to the elliptic curve.

Example 13 (Quadratic forms)

Given an $n \in \mathbb{N}$, can we write n as a sum of k square numbers $a_1^2, \cdots, a_k^2$? Denote the number of ways to do this $S_k(n)$.

For example, $S_2(5) = 8$ (we can use positive and negative numbers, and we can swap their order). It turns out there's a closed form

$$s_2(n) = 2\left(1 + \left(\frac{-1}{n}\right)\right) \sum_{d|n} \left(\frac{-1}{d}\right),$$

but how are we supposed to relate this to modular forms? Well, consider the theta function

$$\Theta(z) = \sum_{j \in \mathbb{Z}} e^{2\pi i z \cdot j^2} = \sum_{j \in \mathbb{Z}} q^{j^2}$$

(where we denote $q = e^{2\pi i z}$). Then this is a generating function, and the q^n coefficient c_n of Θ^k is the number of ways to write n as the sum of k squares. And it turns out that Θ^k is actually a modular form – just not as defined as above. (We'll be more precise in the future.) It has weight $\frac{k}{2}$, which can be pretty interesting to study as well.

And finally, how do we construct L-functions associated to a modular form? For any modular form, $f(z+1) = f(z)$ (because $z + 1$ is the action of $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ on z). So that means we can write f as a Fourier series

$$f(z) = \sum_{n=0}^{\infty} a_n e^{2\pi i n z},$$

and the Fourier coefficients a_n will go into the L-function (except tossing the constant term). This turns out to also have nice properties – this has something to do with "eigenvalues of the Hecke operators." We'll see how everything is connected in the next few months!

2 February 6, 2020

Serre 7.1 – Dhruv Rohatgi

We'll start by talking about the modular group, its action on $\mathbb{H}$, and its fundamental domain.

Definition 14

The **upper half-plane** is defined to be $\mathbb{H} = \{z \in \mathbb{C} : \operatorname{Im} z > 0\}$.

Definition 15

Given any ring R, we can define $SL_2(R)$ to be the multiplicative group of 2×2 matrices $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ with $a, b, c, d \in R$ and $ad - bc = 1$. (We'll be using $R = \mathbb{Z}$ here.)

We can then also define an **action** of $SL_2(\mathbb{Z})$ on the upper half-plane from $SL_2(\mathbb{Z}) \times \mathbb{H}$ to $\mathbb{H}$ via

$$gz = \phi(g, z) = \frac{az + b}{cz + d}, \quad g = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

We need to check that this is a group action first – we just check that $g(hz) = (gh)z$, which we can do with direct calculation. To check that gz is in $\mathbb{H}$ if z is in $\mathbb{H}$, notice that

$$\operatorname{Im}(gz) = \operatorname{Im} \frac{az + b}{cz + d} = \frac{\operatorname{Im}((az + b)(c\overline{z} + d))}{|cz + d|^2}.$$

Expanding the numerator gives $(ad - bc) \operatorname{Im} z = \operatorname{Im} z$ (since $ad - bc = 1$), and then we're dividing by something positive. So this means that gz lies in the upper half-plane.

What can we say about this group action? Notice that $-I$ acts trivially on the upper half-plane (because we have $-z$ divided by -1).

Definition 16

The **projective special linear group** $G = PSL_2(\mathbb{Z})$ is defined as $SL_2(\mathbb{Z})/\{\pm I\}$. This is also called the **modular group**.

We can consider the induced group action of this modular group on $\mathbb{H}$, and a useful thing to have is a set of generators. If we define

$$S = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix},$$

then $Sz = -\frac{1}{z}$ and $Tz = z + 1$.

Theorem 17

S and T generate G.

We can prove this algebraically, but a more geometric approach works well here. Define the region

$$\mathcal{D} = \{z \in \mathbb{H} : -\frac{1}{2} \le \operatorname{Re} z \le \frac{1}{2}, |z| \ge 1\}.$$

Other than the boundary, we'll show that this domain contains one representative from each orbit.

Theorem 18

We have the following facts:

1. There's a surjective map from $\mathcal{D} \to \mathbb{H}/G$. Moreover, for all $z \in \mathbb{H}$, there exists a $g \in \langle S, T\rangle$ such that $gz \in \mathcal{D}$.
2. This map is also "mostly" injective aside from "boundary cases:" if $z, z' \in \mathcal{D}, z \neq z'$, but $z' = gz$ for some $g \in G$, then $\text{Re}(z) = \pm\frac{1}{2}$ and $g = T^{\pm 1}$ or $|z| = 1$ and $g = S$.
3. Finally, every $z \in \mathcal{D}$ has trivial stabilizer except for $i, \rho, -\overline{\rho}$, where ρ is $-\frac{1}{2} + \frac{\sqrt{3}}{2}$.

Why does this imply that S and T generate G? Let z be in the interior of $\mathcal{D}$, and let $g \in G$. By (1), there is some $h \in (S, T)$ such that $h(gz) \in \mathcal{D}$, and by (2) this means $hgz = z$. That means by (3) that $hg = I$, which means $g = h^{-1} \in \langle S, T\rangle$.

Proof. We'll first prove surjectivity: for any $z \in \mathbb{H}$, there exists some n such that $T^n z$ has real part between $-\frac{1}{2}$ and $\frac{1}{2}$. If this point has magnitude 1 or larger, we're done. Otherwise, apply S, which increases the imaginary coordinate, and then apply more Ts to fix the real part again. To prove this terminates, every time we apply an S, we have a sequence of imaginary parts $\text{Im}(z) < \text{Im}(z_2) < \cdots$, where $z_2 = ST^n z$ and so on. But the set

$$\{\text{Im}(gz) : g \in G \wedge \text{Im}(gz) > \text{Im}(z)\}$$

is finite, because $\text{Im}(gz) = \frac{\text{Im}(z)}{|cz+d|^2}$ can only be at least $\text{Im}(z)$ if $|cz+d| \le 1$, which can happen only for a finite number of (c, d).

To show the other parts, take $z, z' \in \mathcal{D}$ so that $z' = gz, g \in G$. Without loss of generality, say that $\text{Im}(z') \ge \text{Im}(z)$, which tells us that $|cz + d| \le 1$ (just like the above part). Since we're in the fundamental domain $\mathcal{D}$, the imaginary part of $\text{Im}(z) \ge \frac{\sqrt{3}}{2}$ and $|cz + d| \ge |c|\,\text{Im}(z)$, which leaves a finite number of details to check. □

Serre 7.2.1 – Vanshika Jain

Definition 19

A function f is **weakly modular** of weight $2k, k \in \mathbb{Z}$ if f is **meromorphic** (analytic everywhere except at poles) on $\mathbb{H}$ and satisfies

$$f(z) = (cz+d)^{-2k} f\left(\frac{az+b}{cz+d}\right)$$

for all $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SL_2(\mathbb{Z})$.

Using the quotient rule (and the fact that $ad - bc = 1$), we find that

$$\frac{d(gz)}{dz} = \frac{1}{(cz+d)^2},$$

so the main relation of our definition can be written as

$$f(gz)d(gz)^k = f(z)dz^k.$$

One way to interpret this is to think of the differential form $f(z)dz^k$ as being invariant under action by G. Since G is generated by S and T, it suffices to check invariance under S and T to check the weakly modular condition.

Proposition 20

Let f be a meromorphic function on $\mathbb{H}$. Then v is weakly modular of weight $2k$ if and only if

$$f(z+1) = f(z), \quad f\left(-\frac{1}{z}\right) = z^{2k} f(z)$$

(corresponding to action by T and S, respectively).

Both of these just come from plugging in the matrices for S and T into the weakly modular condition. This tells us that f is periodic, so we can apply the **change of variables** $q = e^{2\pi iz}$ to get a new function $\tilde{f}$ which is meromorphic on the unit disk with 0 removed. Specifically, we have

$$\tilde{f} = \sum_{-\infty}^{\infty} a_n q^n, \quad f = \sum_{-\infty}^{\infty} a_n e^{2\pi inz}.$$

Notice that as $z \to i\infty$, $q \to 0$.

Definition 21

If $\tilde{f}$ extends to a meromorphic (resp: holomorphic) function at the origin, then f is **meromorphic (resp: holomorphic) at ∞**.

In these cases, the infinite sum $\sum a_n q^n$ only needs to sum from $-m$ to ∞ and 0 to ∞, respectively (corresponding to a pole of order m and a holomorphic function, respectively).

Definition 22

A weakly modular function is **modular** if it is meromorphic at ∞ (that is, $\tilde{f}$ has at most a pole at 0). If f is holomorphic at ∞, we can also define the **value** of f at infinity via $f(\infty) = \tilde{f}(0)$.

So modular functions are analytic on H except at poles, invariant under transformations of G (up to some scaling factors), and $\tilde{f}$'s Laurent expansion has a finite-order pole at 0.

Definition 23

A **modular form** is a modular function that is holomorphic everywhere (including ∞). A modular form which has value 0 at infinity is called a **cusp form**.

So to recap, a modular form of weight $2k$ can be written in the form

$$f(z) = \sum_{n=0}^{\infty} a_n q^n = \sum_{n=0}^{\infty} a_n e^{2\pi inz},$$

converging in the unit disk $|q| < 1$, satisfying $f\left(-\frac{1}{z}\right) = z^{2k} f(z)$ – it's a **cusp form** if $a_0 = 0$.

Example 24

If we're given two modular forms f, f' with weight $2k, 2k'$, then ff' is a modular form of weight $2k + 2k'$ – we can check the two equations (action under T and S) both hold.

Example 25

The function

$$q\prod_{[} n = 1]^{\infty}(1-q^n)^{24} = q - 24q^2 + \cdots$$

is a cusp form of weight 12, and it will come up later on.

Serre 7.2.2 - Swapnil Garg

Definition 26

A **lattice** Γ of a finite-dimensional vector space V (of dimension n) is a subgroup of V isomorphic to $\mathbb{Z}^n$. Γ must also span V, which means that it contains a basis of V.

Example 27 (Non-example)

Taking the set of points $(a + b\sqrt{2}, 0)$ (integer combinations of 1 and $\sqrt{2}$). The result is isomorphic to $\mathbb{Z}^2$ but does not span $\mathbb{R}^2$, so it is not a lattice.

If we look at lattices in $\mathbb{R}^2$, we can generate any lattice with two basis vectors (w_1, w_2). To avoid double-counting, we'll be a bit more precise:

Definition 28

Let $\mathcal{R}$ be the set of lattices in $\mathbb{C}$ if we look at it as $\mathbb{R}^2$, and let M be the set of ordered pairs (w_1, w_2) with $w_1, w_2 \in \mathbb{C} \setminus 0$ and $\frac{w_1}{w_2} \in \mathbb{H}$.

Observe that all lattices can then be generated by (w_1, w_2), so we have a surjective map from M to $\mathcal{R}$. A natural next question to ask is when two ordered pairs generate the same lattice – this is where $SL_2(\mathbb{Z})$ comes in. First, we'll need a lemma:

Lemma 29

Suppose that $v_1 = gv_2$, where $v_1 = (w_1, w_2), v_2 = (w_1', w_2')$ are two-dimensional (complex) vectors and g is a 2×2 matrix. If g has positive determinant, then $z = \frac{w_1}{w_2}$ and $z' = \frac{w_1'}{w_2'}$ have the same sign if and only if g has positive determinant.

(This was basically proven by Dhruv's part of the lecture.)

Proposition 30

$v_1, v_2 \in M$ map to the same lattice in $\mathcal{R}$ if and only if $v_1 = gv_2$ for some $g \in SL_2(\mathbb{Z})$.

Proof. The backwards direction is clear, because $v_1 = gv_2$ is a lattice transformation – we map our basis vectors to other vectors in our lattice, and we have the same unit cell size.

The forwards direction is very similar: if $v_1 = (w_1, w_2)$ and $v_2 = (w_1', w_2')$, then $v_1 = aw_1 + bw_2$ and $v_2 = cw_1 + dw_2$. This means that $g = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, and having the same unit cell cize requires the determinant to be ± 1. By the above lemma, we know that the determinant of g is positive, so $g \in SL_2(\mathbb{Z})$. □

This means that $\mathcal{R}$ is isomorphic to $M/SL_2(\mathbb{Z})$. Now we also want to mod out by the action of $\mathbb{C}^*$ by scalar multiplication (which scale, stretch, and rotate our lattices). If we consider the map from M to $\mathbb{H}$ sending (w_1, w_2) to $\frac{w_1}{w_2}$, notice that (w_1, w_2) maps to the same point as $(\lambda w_1, \lambda w_2)$. So $\mathbb{H}$ is bijective with $M/\mathbb{C}^*$, meaning pairs of basis vectors mod $\mathbb{C}^*$ are identified with the upper half-plane. And if we mod this out by $SL_2(\mathbb{Z})$, we find that $\mathcal{R}/\mathbb{C}^*$ is isomorphic to $\mathbb{H}/G$, where G is the modular group defined earlier. (It's okay to mod out by G rather than $SL_2(\mathbb{Z})$ because ± 1 don't do anything.) And this is why we introduce lattice – $\mathcal{R}/\mathbb{C}^*$ is very closely related to $\mathbb{H}/G$, which is closely related to the fundamental domain $\mathcal{D}$.

Definition 31

A **lattice function** $F : \mathcal{R} \to \mathbb{C}$ has weight **2k** if $F(\lambda\Gamma) = \lambda^{-2k}F(\Gamma)$ for all $\lambda \in \mathbb{C}, \Gamma \in \mathcal{R}$.

We can think of F as acting on the basis vectors instead of the whole lattice – this gives us a map $F : M \to \mathbb{C}$, where $F(w_1, w_2)$ is F of the lattice generated by (w_1, w_2). The lattice function condition then becomes

$$F(\lambda w_1, \lambda w_2) = \lambda^{-2k}F(w_1, w_2),$$

so $w_2^{2k}F(w_1, w_2)$ is invariant. This motivates us to define

$$f\left(\frac{w_1}{w_2}\right) = w_2^{2k}F(w_1, w_2).$$

Also, because F is invariant under $SL_2(\mathbb{Z})$, we know that

$$f(z) = F(z, 1) = F(az + b, cz + d) = (cz + d)^{-2k}f\left(\frac{az + b}{cz + d}\right),$$

which is exactly the modular function condition. So adding a few extra conditions means that f can be a modular function!

3 February 11, 2020

Serre 7.2.3 – Christian Altamirano

Last week, we defined modular functions, and we'll be presenting some examples today. We'll start with a lemma:

Lemma 32

Let Γ be a lattice in $\mathbb{C}$. Then $\sum'_{\gamma \in \Gamma} \frac{1}{|\gamma|^\sigma}$ is convergent for $\theta > 2$, where $\sum'$ denotes a sum over all nonzero elements of the lattice.

Proof. There are two ideas we can use here. First, we can majorize the series under (a constant times) the double integral

$$\iint \frac{dxdy}{(x^2 + y^2)^{\theta/2}}$$

by writing out the double integral as a Riemann sum. Then the double integral is easy to evaluate by using polar coordinates.

Another idea is to bound the points in the lattice with $|\gamma|$ between n and $n + 1$ – there are $O(n)$ of these by an area argument because $O([n + 1]^2 - n^2) = O(n)$, so the infinite sum for the lattice is convergent if and only if $\frac{1}{n^{\sigma-1}}$ is convergent because

$$\sum_{\gamma\in\Gamma} \frac{1}{|\gamma|^\sigma} \le \sum_{n=1}^{\infty} O(n)\frac{1}{n^\theta} = K\sum_{n=1}^{\infty}\frac{1}{n^{\sigma-1}}.$$

□

With this, we can construct an example:

Example 33

Let $k > 1$ be an integer and let Γ be a lattice. Then define a function on lattices

$$G_k(\Gamma) = \sum_{\gamma\in\Gamma}{}' \frac{1}{\gamma^{2k}}.$$

From the above lemma, this is an absolutely convergent sum – this is called the **Eisenstein series of index *k***. We know that our lattice Γ can be defined by two complex numbers w_1, w_2, and any point can be written as $mw_1 + nw_2$ (an integer linear combination), so we can instead define

$$G_k(w_1, w_2) = \sum_{(m,n)}{}' \frac{1}{(mw_1 + nw_2)^{2k}}.$$

Note that the "shape" of the lattice depends mostly on $z = \frac{w_1}{w_2}$, we can define

$$G_k(z) = G_k(z, 1) = \sum_{(m,n)}{}' \frac{1}{(mz+n)^{2k}}.$$

Proposition 34

$G_k(z)$ is a modular form of weight $2k$ for all integers $k > 1$, and we have $G_k(\infty) = 2\zeta(2k)$.

Proof. First, we show that $G_k(z)$ is weakly modular (of weight $2k$). We can show this by looking at the two transformations under T and S: recall that f is weakly modular of weight $2k$ if it satisfies

$$f(z) = f(z+1), \quad f\left(\frac{1}{z}\right) = z^{-2k} f(z),$$

and we can check that both of these equations hold.

The next step is to show $G_k(z)$ is holomorphic everywhere (including at ∞). Let $z \in D$ (the fundamental domain of the modular group) – note that

$$|mz+n|^2 \ge m^2 - mn + n^2 = |m\rho - n|^2,$$

where $\rho = e^{2\pi i/3}$, by a direct calculation, and thus

$$G_k(z) \le \sum \frac{1}{(m\rho - n)^{2k}}$$

is convergent by Lemma 32 – in fact, it converges **normally**. Now, given any $z \in D$ and any $g \in G = \langle S, T\rangle$, we get $G_k(gz)$, which will also converge normally (because we're transforming the domain D). Thus G_k is holomorphic in the upper half-plane.

Finally, we need to show that G_k is holomorphic at ∞, which is the same as showing that $G_k(z)$ has a limit as $z \to \infty$. We can assume that z is in the fundamental domain (meaning we take $z \to i\infty$), and then

$$\lim_{z \to \infty} \sum_{m,n}{}' \frac{1}{(mz+n)^{2k}} = \lim_{z \to \infty} \sum_{n \in \mathbb{Z}}{}' \frac{1}{n^{2k}} = 2\zeta(2k)$$

because any term with a z will disappear (the denominator becomes infinity). □

The main idea of **normal convergence** is that a function

$$f = \sum_n f_n$$

converges normally if $\sum_n \max |f_n(z)|$ converges. Basically, it's a very strong condition – it implies that the series converges at any point.

Serre 7.3.1 – Anton Trygub

Definition 35

Let f be a meromorphic function on $\mathbb{H}$. Then the **order of f at p** is the integer n so that $\frac{f}{(z-p)^n}$ is a holomorphic function at p and $f(p) \neq 0$. If n is positive, then p is called a **zero** of f, and if n is negative, then p is called a **pole** of f. Denote this order $\nu_p(f)$.

We know that by definition, a modular function f satsifies

$$f(z) = (cz+d)^{-2k} f\left(\frac{az+b}{cz+d}\right)$$

for all $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SL_2(\mathbb{Z})$. This tells us that

$$\nu_z(f) = \nu_{\frac{az+b}{cz+d}}(f);$$

in other words, we have $\nu_p(f) = \nu_{gp}(f)$ for any $p \in \mathbb{H}$ and $g \in G$. So this function ν is constant on every orbit of p – it only depends on the image of p in $\mathbb{H}/G$.

For convenience, we'll also define the order $v_\infty(f)$ to be the order of $q = 0$ for $\tilde{f}(q)$, where the function $\tilde{f}$ is defined so that $q = e^{2\pi i z}$ and $\tilde{f}(q) = f(z)$.

Claim 36. *Let f be a modular function of weight $2k$. Then f has only finitely many zeros and poles in the fundamental domain $D = \{|\operatorname{Re}(z)| \le \frac{1}{2}, \quad |z| \ge 1, \operatorname{Im}(z) > 0\}$.*

Proof. Since $\tilde{f}$ is meromorphic at $q = 0$, and zeros and poles are isolated for meromorphic functions, there exists a neighborhood of 0 where $\tilde{f}$ has no zeros and poles – specifically, there is an $r > 0$ such that $\tilde{f}$ has no zeros for all $0 < |q| < r$. Because $\tilde{f}(e^{2\pi i z}) = f(z)$, this means that $f(z)$ does not have any zeros or poles for $\operatorname{Im}(z) > \frac{\log(1/r)}{2\pi}$. So any zeros or poles of f in the fundamental domain are in the compact region

$$D_r = \{x \in D : \operatorname{Im}(x) \le \frac{\log(1/r)}{2\pi}\}$$

and there can only be a finite number of zeros or poles here, as desired. □

Definition 37

Let e_p be the order of the stabilizer of the point p under G.

From the first theorem of the chapter (where we said that we have an injective function except with a few exceptions), we know that

$$e_p = \begin{cases} 2 & p = g(i) \\ 3 & p = g(\rho) \\ 1 & \text{otherwise} \end{cases}$$

(in each case, for some $g \in G$).

Theorem 38

Let f be a modular function of weight $2k$. Then

$$\nu_\infty + \sum_{p \in \mathbb{H}/G} \frac{1}{e_p} \nu_p(f) = \frac{k}{6}.$$

Our previous claim tells us that there are only finitely many points $p \in \mathbb{H}/G$ with order different from zero, so this definition makes sense. We can also rewrite this with our calculation of e_p:

$$\boxed{\nu_\infty + \frac{1}{2}\nu_i(f) + \nu_\rho(f) + \sum_{\substack{\rho \in \mathbb{H}/G \\ \rho \neq gi, g\rho}} \nu_p(f) = \frac{k}{6}}.$$

Serre 7.3.1 continued – Zack Chroman

We'll now prove the above theorem. We'll denote the sum over p over the domain D which are not i or ρ with the symbol $\sum^*$.

The big theorem we'll need to show this is the following:

Theorem 39

If we integrate a function f around a closed domain A, then

$$\frac{1}{2\pi i} \int_{\delta A} f(z) dz = \sum_{p \in A} \text{Res}_p(f),$$

where the **residue** $\text{Res}_p(f)$ is defined to be the coefficient a_{-1} in the Laurent series expansion

$$f(z) = \sum_{n=-\infty}^{\infty} a_n (z-p)^n.$$

If we apply the residue theorem to the function $\frac{f'}{f} dz$, we get the following nice corollary:

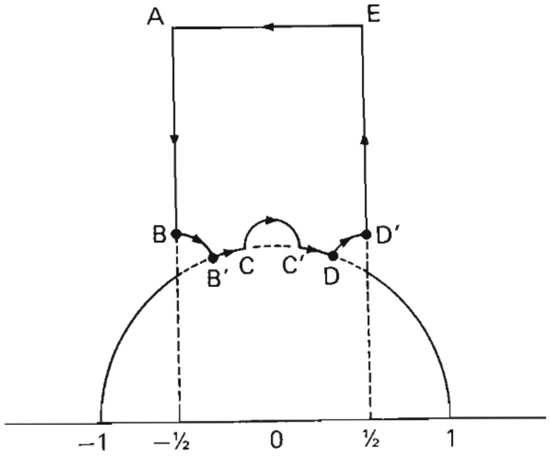
Corollary 40 (Argument principle)

We have

$$\frac{1}{2\pi i} \int_{\delta A} \frac{f'}{f} dz = \sum_{p \in A} \nu_p(f).$$

We'll skip the proof of this for now – it is an exercise in expanding out the Laurent series.

To prove the theorem we want, we will use the residue theorem to evaluate a certain integral: we take the fundamental domain D, but we avoid a neighborhood of the points ρ, i, and $\rho + 1$ (picture taken from Serre):



Here, we're picking A and E to be high enough so that all of the zeros and poles in the fundamental domain are below line AE, and we're also making the radius of the curves from B to B', C to C', and D to D' go to 0.

We assume that there's no poles along the parts from A to B or E to D' – if there are, then we just avoid those by drawing an identical curve on the left and right.

Call our curve $\mathcal{C}$. By the residue theorem, we know that

$$\int_{\mathcal{C}} \frac{f'}{f} dz = 2\pi i \sum^{*} \nu_p(f),$$

but we can also break this up along each of the individual parts. The integral along A to B and D' to E are identical by periodicity, but we're going in opposite directions, so they cancel out.

Now, under the transformation $z \mapsto e^{2\pi i z}$, the line AE gets mapped to a circle ω around 0, and thus

$$\int_E^A \frac{f'}{f} dz = \int_\omega \frac{df}{f} = -2\pi i \nu_\infty(f).$$

(The negative sign comes from the orientation of the circle.) The rest of the curves, like from B to B', are all partial circles. To evaluate these, we need a lemma:

Lemma 41

Integrating along a small circular arc around a pole p of angle α yields $\text{Res}_p(f) \cdot \alpha i$.

The idea is that the integral all the way around the pole is $2\pi i \text{Res}_p(f)$, so it's just proportional with the arc length.

Proof. First of all, we can ignore all power series terms except the $\frac{a_{-1}}{z-p}$ term (everything else goes away because it integrates to 0). Thus, we want the integral

$$\int_{\text{arc}} \frac{a_{-1}}{z-p} dz.$$

Substituting $z = p + e^{2\pi i \theta}$, we have $dz = 2\pi i e^{2\pi i \theta} d\theta$ and thus our integral becomes

$$\int_0^x \frac{a_{-1}}{e^{2\pi i \theta]}} \cdot 2\pi i e^{2pii\theta} d\theta = xa^{-1} \cdot 2\pi i,$$

as desired. □

With this, note that as $r \to 0$, the arc from B to B' becomes an arc of angle $\frac{\pi}{3}$ (because the angle between the vertical line and the tangent at ρ is $\frac{\pi}{3}$). This means that

$$\int_B^{B'} \frac{f'}{f} dz = -\frac{2\pi i}{6} \nu_\rho(f),$$

and similarly

$$\int_D^{D'} \frac{f'}{f} dz = -\frac{2\pi i}{6} \nu_\rho(f).$$

Meanwhile, the circular arc from C to C' approaches a semicircle, so

$$\int_C^{C'} \frac{f'}{f} dz = -\frac{2\pi i}{2} \nu_i(f).$$

Finally, note that the matrix $S = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ sends the arc $B'C$ to the arc DC' (oriented in that direction). The definition of a modular function tells us that

$$f(Sz) = z^{2k} f(z) \implies \int_{B'}^{C} \frac{df(Sz)}{f(Sz)} dz = \int_{B'}^{C} 2k\frac{dz}{z} + \frac{df}{f} = \int_{D}^{C'} \frac{df}{f}.$$

This means that the sum

$$\int_{B'}^{C} \frac{df}{f} + \int_{C'}^{D} \frac{df}{f} = \int_{B'}^{C} -2k\frac{dz}{z}$$

(where the signs come from us switching the order of integration), and now this is just the residue theorem on the unit circle. The angle from B' to C goes to $\frac{\pi}{12}$, so this evaluates to $-2k \cdot -\frac{2\pi i}{12} = \frac{2\pi i k}{6}$.

Putting everything together, this means that

$$\int_C \frac{f'}{f} dz = -2\pi i \left(\nu_\infty(f) + \frac{\nu_\rho(f)}{3} + \frac{\nu_i(f)}{2} - \frac{k}{6} \right),$$

but we know from before that this integral is also equal to $2\pi i \sum^* \nu_p(f)$. Rearranging and canceling the $2\pi i$ gives us the result we want.

4 February 13, 2020

Serre 7.3.2 – David Wu

We've been considering properties of modular forms, and now we will look at the whole space of modular forms together. Let $\mathcal{M}_k$ be the space of modular forms of weight $2k$, and let $\mathcal{M}_k^0$ be the space of cusp forms. Both are vector spaces over $\mathbb{C}$, because the sum of holomorphic functions is holomorphic, and 0 is always a modular form of weight k.

Here's how we can relate modular forms and cusp forms:

Proposition 42

We have $\mathcal{M}_k = \mathcal{M}_k^0 \oplus \mathbb{C}G_k$ (which means we add a complex number times the Eisenstein series).

Proof. Consider the map from M_k to $\mathbb{C}$ sending f to $f(\infty)$. By definition of cusp forms, the kernel of this map is M_k^0, and the image is a subspace of $\mathbb{C}$ (so either all of $\mathbb{C}$ or just 0). But we know that $G_k(\infty) \neq 0$, so we can indeed decompose as stated. □

The nice thing about Eisenstein series is that we can work directly with them – in contrast, we don't actually have that many examples of cusp forms. But one good example is

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

where $g_2 = 60G_2$ and $g_3 = 140G_3$. (This is an element of M_6^0 – a cusp form of weight 12.)

Theorem 43

The space of modular forms $\mathcal{M}_k$ is trivial for $k < 0$ and $k = 1$. For $k = 0, 2, 3, 4, 5$, $\mathcal{M}_k$ is a vector space spanned by $1, G_2, G_3, G_4, G_5$ respectively, and M_k^0 is trivial. Also, if we multiply by Δ, then we have an isomorphism between $\mathcal{M}_{k-6}$ and $\mathcal{M}_k^0$.

Basically, we can classify for small values of k and then move to higher weights as well.

Proof. Recall from last time the formula

$$v_\infty(f) + \frac{1}{2}v_i(f) + \frac{1}{3}v_\rho(f) + \sum_{H/G}^{*} v_p(f) = \frac{k}{6}.$$

All of these orders are nonnegative because we're working with holomorphic functions f, so we must have $k \ge 0$ to have a nonzero function f. Also, if $k = 1$, then there are no terms that can give a contribution of $\frac{1}{6}$ – the smallest contribution is $\frac{1}{3}$.

Also, if we have a cusp form, then $v_\infty(f) \ge 1$ by definition, so the left hand side is at least 1 (and therefore for $k < 6$, there are no cusp forms other than 0). By Proposition 42, this means $M_k = \mathbb{C}G_k$, so M_k is indeed just a one-dimensional vector space.

Finally, let's first prove a few properties about our discriminant function Δ. If we apply our counting formula above to G_2, we find that $v_\rho(G_2) = 1$ and $v_p(G_2) = 0$ for all other points p. Similarly, $v_i(G_3) = 1$ and $v_p(G_3) = 0$ for all other p. This means that if we look at our discriminant $\Delta = g_2^3 - 27g_3^2$, it will be nonzero at i (because there is a zero for the g_3 term, but not for the g_2 term). Therefore, Δ is not identically zero, and because it is a cusp form, $v_\infty(\Delta) \ge 1$. Again applying our counting formula, all of the other v_ps must be zero for Δ, so Δ is nonzero on all of $\mathbb{H}$.

So now take any $f \in M_k^0$, and define $g = \frac{f}{\Delta}$ – this is holomorphic because Δ is zero everywhere except at ∞ (where f already has a zero), and that means we can find the orders of the zeros of g:

$$v_p(g) = \begin{cases} v_p(f) & p \ne \infty \\ v_p(f) - 1 & p = \infty \end{cases}$$

and computing again with the counting formula shows that we do indeed end up in $\mathcal{M}_{k-6}$. □

Corollary 44

We have an explicit formula for all $k \ge 0$:

$$\dim \mathcal{M}_k = \begin{cases} \left\lfloor \frac{k}{6} \right\rfloor & k \equiv 1 \bmod 6 \\ \left\lfloor \frac{k}{6} \right\rfloor + 1 & \text{otherwise.} \end{cases}$$

This tells us that these vector spaces are finite-dimensional, which is nice!

Proof. Applying Theorem 43 gives the desired result for $k = 0, 1, 2, 3, 4, 5$ (verify directly). To pass to larger k, note that adding 6 to k makes the right-hand side of the equation just increase by 1, and by Proposition 42,

$$\mathcal{M}_{k+6} = \mathcal{M}^0_{k+6} \oplus \mathbb{C}G_{k+6} = \mathcal{M}_k \oplus \mathbb{C}G_{k+6}$$

(last equality from Theorem 43), which means the dimension of $\mathcal{M}_{k+6}$ just tacks on a one-dimensional subspace. □

Corollary 45

$\mathcal{M}_k$ has a basis $\{G_2^m G_3^n, m, n \in \mathbb{Z}_{\geq 0}, 2m + 3n = k\}$.

Proof. First we show that this set of vectors generate the vector space, and then we show that they are linearly independent. If k is small, specifically for all $k \leq 3$, we can check directly that M_2 is spanned by G_2 and M_3 is spanned by G_3.

For all $k > 3$, we again use induction: the Chicken McNugget Theorem tells us that there exist α, β with $2\alpha + 3\beta = k$. So if we compute the weight of $g = G_2^\alpha G_3^\beta$, the weight of the modular form will be k, and it is nonzero at ∞ (because G_2 and G_3 are nonzero at ∞). So for any modular form $f \in \mathcal{M}_k$, there exists $\lambda \in \mathbb{C}$ such that $f - \lambda g$ is a cusp form, which means

$$f - \lambda g = \Delta h, \quad h \in \mathcal{M}_{k-6}.$$

By the inductive hypotehsis, this means we can write

$$f - \lambda g = (g_2^3 - 27g_3^2)(G_2^a G_3^b)$$

for some a, b, and expanding this out gives the appropriate result.

To show that these monomials are independent, say that we have a dependence relation

$$a_1 G_2^\alpha G_3^\beta + a_2 G_2^{\alpha-3} G_3^{\beta+2} + \cdots = 0.$$

Dividing through by a power of G_3^β tells us that $\frac{G_2^3}{G_3^2}$ satisfies a polynomial equation, so the fundamental theorem of algebra tells us that $\frac{G_2^3}{G_3^2}$ is completely constant, which is not true (because G_2 is zero at ρ while G_3 is not, and this is not the zero function). □

So we've shown that we can classify the vector spaces of modular forms, and we have explicit formulas for the basis elements and dimensions of these spaces

Serre 7.3.3 – Shreyas Balaji

We'll talk about a specific nice modular function:

Definition 46

The **modular invariant** is given by

$$j = \frac{1728 g_2^3}{\Delta}.$$

To motivate the 1728 coefficient in the numerator, let's think about the series expansion: one way we can make it nice is to make the residue at ∞ equal to 1 (because j has a simple pole at ∞, which we'll show).

Proposition 47

j is a modular function of weight 0.

Proof. g_2 is a modular form of weight $2k = 4$, so g_2^3 is a modular form of weight 12. Δ also has weight 12, so the quotient has weight zero. □

Proposition 48

j is holomorphic on $\mathbb{H}$, and it has a simple pole at ∞.

Proof. Recall from the previous section that g_2 is holomorphic on $\mathbb{H}$ and at ∞, and Δ has a simple zero at ∞ and is nonzero everywhere else. Dividing will not introduce any zeros or poles except at ∞. □

Proposition 49

j defines by **passage to quotient** a bijection between $\mathbb{H}/G \to \mathbb{C}$.

Proof. Here, "passage to quotient" means we quotient everything in the upper half plane by G – for this to be valid, j should be equal on the equivalence classes in $\mathbb{H}/G$, but remember that j has weight zero (so this is well-defined). To show that we have a bijection, we need to show that for any $\lambda \in \mathbb{C}$, there exists a unique $\omega \in \mathbb{H}/G$ such that $j(\omega) = \lambda$.

Define the function $f_\lambda : \mathbb{H}/G \to \mathbb{C}$ via

$$f_\lambda = 1728 g_2^3 - \lambda\Delta.$$

We're trying to show that λ has a unique zero: we can apply the counting zeros formula to find that

$$v_\infty(f) + \frac{1}{2}v_i(f) + \frac{1}{3}v_\rho(f) + \sum_{H/G}^{*} v_p(f) = \frac{k}{6},$$

where $k = 6$ in this case. There are a few options: we can have a zero at ∞, two zeros at i, 3 zeros at ρ, or one zero at a certain point. But there's always exactly one point at which we have a zero, which is what we want. □

Proposition 50

Let f be a meromorphic function over $\mathbb{H}$. Then the following are equivalent:

1. f is a modular function of weight 0.
2. f is a quotient of two modular forms of the same weight.
3. f is a rational function of j.

Proof. (3) implies (2) implies (1), so all we need to do is to show that (1) implies (3). Let f be a modular function of weight 0 – in fact, we can assume that f is holomorphic over $\mathbb{H}$, because it starts off meromorphic with finitely many poles and then we can multiply by some polynomial in j to get rid of those.

The discriminant Δ is zero at ∞, so the function $g = \Delta^n f$ is holomorphic at ∞ for some n. Now g has weight $12n$ (because f has weight 0 and Δ has weight 12), so we know that g is a linear combination

$$g = \sum_{2\alpha+3\beta=6n} G_2^\alpha G_3^\beta.$$

It suffices to show that each term is a rational function of j by linearity. We have

$$f = \frac{G_2^\alpha G_3^\beta}{\Delta^n},$$

and we know that $2\alpha + 3\beta = 6n$, so α is a multiple of 3 and β is a multiple of 2. Let $\alpha = 3p$ and $\beta = 2q$, and now we can write

$$f = \frac{G_2^{3p} G_3^{2q}}{\Delta^{p+q}}$$

(the denominator $p+q$ comes from making sure f has weight 0). And now f must be a rational function of j: $\frac{G_2^{3p}}{\Delta^p}$ and $\frac{G_3^{2q}}{\Delta^q}$ are both functions of j, and we're done. □

Serre 7.4.1 – Andrew Gu

We'll be talking about Bernoulli numbers, which will be helpful for computing some coefficients.

Definition 51

The **Bernoulli numbers** are defined by the power series expansion

$$\frac{x}{e^x - 1} = 1 - \frac{x}{2} + \sum_{k=1}^{\infty} (-1)^{k+1} B_k \frac{x^{2k}}{(2k)!}.$$

The infinite sum has no odd-degree terms; to check this, note that

$$\frac{x}{e^x - 1} + \frac{x}{2} = \frac{x}{2}\left(\frac{e^x + 1}{e^x - 1}\right),$$

which is an even function. As an example, the first few Bernoulli numbers are

$$B_1 = \frac{1}{6}, B_2 = \frac{1}{30}, B_3 = \frac{1}{42}, \cdots.$$

The Bernoulli numbers are sometimes defined differently via

$$\frac{x}{e^x - 1} = \sum_{k=0}^{\infty} \frac{b_k x^k}{k!},$$

so we don't throw away the odd-degree terms. These give the same numbers up to some small changes, but we'll just use the B_k (because that's what the book uses).

Theorem 52

We have a formula for the zeta function at even values: for all integers $k \geq 1$,

$$\zeta(2k) = \frac{2^{2k-1}}{(2k)!} B_k \pi^{2k}.$$

Proof. Set $x = 2iz$ in the definition of the Bernoulli numbers to find that

$$\frac{2iz}{e^{2iz} - 1} = 1 - iz - \sum_{k=1}^{\infty} \frac{2^{2k} B_k z^{2k}}{(2k)!}.$$

Moving the iz term to the left gives

$$\boxed{1 - \sum_{k=1}^{\infty} \frac{2^{2k}}{(2k)!} B_k z^{2k}} = z \cot z$$

by expanding out the definitions of the exponentials. But we can also find a different formula for $z \cot z$: starting with the Euler product

$$\sin z = z \prod_{n=1}^{\infty} \left(1 - \frac{z^2}{n^2\pi^2}\right).$$

We apply logarithmic differentiation, meaning we send $f \to \frac{f'}{f}$. This turns products into sums:

$$\frac{\cos z}{\sin z} = \frac{1}{z} + \sum_{n=1}^{\infty} \frac{-2z/(n^2\pi^2)}{1 - z^2/(n^2\pi^2)}.$$

Multiplying both sides by z yields

$$z \cot z = 1 - 2\sum_{n=1}^{\infty} \frac{z}{n^2\pi^2 - z^2}.$$

This is almost a power series expansion – we'll now expand this fraction as a power series, so

$$z \cot z = 1 - 2\sum_{n=1}^{\infty} \frac{z^2/(n^2\pi^2)}{1 - z^2/(n^2\pi^2)} = 1 - 2\sum_{n=1}^{\infty}\sum_{k=1}^{\infty} \left(\frac{z^2}{n^2\pi^2}\right)^k$$

is valid as long as $|z| < \pi$. Exchanging the sums yields

$$\boxed{1 - 2\sum_{k=1}^{\infty} \frac{z^{2k}}{\pi^{2k}} \sum_{n=1}^{\infty} \frac{1}{n^{2k}}}.$$

Comparing coefficients with our two boxed equations yields exactly what we want (we are comparing power series around a neighborhood of 0, so they must agree). □

5 February 13, 2020

Serre 7.4.2 – Michelle Xu

We'll be discussing the Eisenstein series of index k (an example of a modular form)

$$G_k(z) = \sum_{m,n}{}' \frac{1}{(nz+m)^{2k}}$$

We'll be trying to express this as a Taylor expansion in terms of $q = e^{2i\pi z}$.

Lemma 53

For all $k \geq 2$, we have

$$\sum_{m\in\mathbb{Z}} \frac{1}{(z+m)^k} = \frac{1}{(k-1)!}(-2\pi i)^k \sum_{n=1}^{\infty} n^{k-1}q^n.$$

Proof. Recall Euler's sine product formula: taking the log derivative of the expression for $\sin z$ yields

$$x \cot x = 1 - 2\sum_{m=1}^{\infty} \frac{x^2}{m^2\pi^2 - x^2}.$$

Set $x = \pi z$ and dividie both sides by z to yield the expression

$$\pi \cot \pi z = \frac{1}{z} - 2\sum_{m=1}^{\infty} \frac{z}{m^2 - z^2}$$

and we can break up the fraction into a simpler sum

$$= \frac{1}{z} + \sum_{m=1}^{\infty} \frac{1}{z+m} + \frac{1}{z-m} = \sum_{m \in \mathbb{Z}} \frac{1}{z+m}.$$

But we can rewrite the left hand side another way:

$$\pi \cot \pi z = \pi \frac{\cos \pi z}{\sin \pi z} = i\pi \frac{q+1}{q-1},$$

because

$$\frac{q+1}{q-1} = \frac{2\cos^2 \pi z + 2i \sin \pi z \cos \pi z}{-2\sin^2 \pi z + 2i \sin \pi z \cos \pi z} = -i\frac{\cos \pi z}{\sin \pi z},$$

and we can continue to simplify this via

$$i\pi \frac{q+1}{q-1} = i\pi \left(1 + \frac{2}{q-1}\right) = i\pi - 2i\pi \sum_{n=0}^{\infty} q^n.$$

Setting these equal to each other yields

$$\sum_{m \in \mathbb{Z}} \frac{1}{z+m} = i\pi - 2i\pi \sum_{n=0}^{\infty} q^n.$$

Differentiating $k-1$ times (this is why we need $k \geq 2$) yields the result. □

This allows us to get to our Taylor expansion:

Proposition 54

For all $k \geq 2$, we have

$$G_k(z) = 2\zeta(2k) + 2\frac{(2i\pi)^{2k}}{(2k-1)!}\sum_{n=1}^{\infty} \sigma_{2k-1}(n) q^n,$$

where $\sigma_k(n)$ is defined to be the sum of the kth powers of divisors of d: $\sigma_k(n) = \sum_{d|n} d^k$.

Proof. We break up the sum into the contribution from $n = 0$ and $n \neq 0$:

$$G_k(z) = \sum_{m \in \mathbb{Z}}{}' \frac{1}{m^{2k}} + \sum_{n \neq 0} \sum_{m \in \mathbb{Z}} \frac{1}{(nz+m)^{2k}}.$$

The first sum just gives us twice $\zeta(2k)$ by similar arguments to what we've done in the past, and we can pull out a factor of 2 in the other sum as well:

$$= 2\zeta(2k) + 2\sum_{n=1}^{\infty} \sum_{m \in \mathbb{Z}} \frac{1}{(nz+m)^{2k}}.$$

This is what we have in the lemma, except that we replace z with nz and k with $2k$. So we can rewrite our lemma as

$$\sum_{m\in\mathbb{Z}} \frac{1}{(nz+m)^{2k}} = \frac{1}{(2k-1)!}(-2\pi i)^{2k} \sum_{a=1}^{\infty} a^{2k-1} q^{na}.$$

Plugging this into our Eisenstein series yields (relabeling n as d)

$$G_k(z) = 2\zeta(2k) + \frac{2(-2i\pi)^{2k}}{(2k-1)!} \sum_{b=1}^{\infty}\sum_{a=1}^{\infty} a^{2k-1} q^{da}.$$

Then we get a contribution to a_n of a^{2k-1} for every $d|n$k, which gives us what we want. □

Corollary 55

Let $E_k(z) = 1 + \gamma_k \sum_{n=1}^{\infty} \sigma_{2k-1}(n) q^n$, where $\gamma_k = (-1)^k \frac{4k}{B_k}$. Then we can rewrite

$$G_k(z) = 2\zeta(2k) E_k(z).$$

Proof. Last time, we found that for all $k \geq 1$,

$$\zeta(2k) = \frac{2^{2k-1}}{(2k)!} B_k \pi^{2k}.$$

This means that

$$\gamma_k = (-1)^k \frac{4k(2\pi)^{2k}}{2(2k)!\zeta(2k)} = \frac{(2i\pi)^{2k}}{(2k-1)!\zeta(2k)},$$

and this gives us the constant that we want. □

We can see some examples of how this looks for different k.

Example 56

$B_2 = \frac{1}{30}$, so we can write

$$E_2(z) = 1 + 240k \sum_{n=1}^{\infty} \sigma_3(n) q^n.$$

We also know that $\zeta(4) = \frac{\pi^4}{90}$, so we can also write

$$g_2 = 60G_2 = 120\zeta(4)E_2 = \frac{4\pi^4}{3} E_2.$$

In general, every E_k is a polynomial in E_2 and E_3 – this is the same argument as with the G_ks. For example, because $E_4 = E_2^2$ and $E_5 = E_2 E_3$, we can get identities like

$$\sigma_7(n) = \sigma_3(n) + 120 \sum_{m=1}^{n-1} \sigma_3(m)\sigma_3(n-m).$$

Serre 7.4.3 – Andrew Lin

Recall that a modular form is holomorphic everywhere in $\mathbb{H}$, so we can express it as a Fourier series

$$f(z) = \tilde{f}(q) = \sum_{n=0}^{\infty} a_n q^n$$

where $q = e^{2\pi i z}$. A natural question to ask about is the order of growth of the a_n (to understand the contribution of higher-order terms). To answer this, recall that the space of modular forms of weight $2k$ can be decomposed via

$$\mathcal{M}_k = \mathcal{M}_k^0 \oplus \mathbb{C}G_k,$$

where $\mathcal{M}_k^0$ is the cusp forms, and G_k is the Eisenstein series of weight $2k$. We'll study these separately:

Proposition 57

Let $k \geq 2$. For the modular form $f = G_k$, we have $|a_n| = \Theta(n^{2k-1})$. In other words, there exist $A, B > 0$ such that (for all $n \geq 1$) we have

$$An^{2k-1} \leq |a_n| \leq Bn^{2k-1}.$$

Proof. In the previous section, we showed that the coefficients a_n (for $n \geq 1$) satisfied

$$a_n = (-1)^k C\sigma_{2k-1}(n),$$

where C is a constant (depending on k) independent of n, and $\sigma_{2k-1}(n)$ is the sum of the $(2k-1)$th powers of divisors of n. We can bound these magnitudes both from above and below:

$$|a_n| = C\sigma_{2k-1}(n) \geq Cn^{2k-1},$$

and also

$$\frac{|a_n|}{n^{2k-1}} = C\sum_{d|n} \frac{1}{d^{2k-1}} \leq C\sum_{d\geq 1} \frac{1}{d^{2k-1}} = D < \infty$$

(in the first equality we use the fact that summing $d^{2k-1}/n^{2k-1} = 1/(n/d)^{2k-1}$ over all divisors d is the same if we replace n/d with d). This means we have shown that $Cn^{2k-1} < |a_n| < Dn^{2k-1}$, and thus $|a_n| = \Theta(n^{2k-1})$, as desired. □

The subspace of cusp forms is a bit more tricky:

Theorem 58 (Hecke)

For any cusp form f of weight $2k$, we have $|a_n| = O(n^k)$; that is, $\frac{|a_n|}{n^k}$ is bounded as $n \to \infty$.

Proof. By definition, the power series expansion of f has $a_0 = 0$, which means that as $q \to 0$ (meaning $z = x + iy$ tends to $i\infty$), the magnitude of f is $O(q)$. In other words,

$$|f(z)| = O(e^{2\pi i(x+iy)}) = O(e^{-2\pi y}).$$

To relate the magnitude of f to y more explicitly, note that for modular forms f of weight $2k$, the function $\phi(z) = |f(z)|y^k$ is **invariant under the modular group** G. This is because $\text{Im}(gz) = \frac{\text{Im}(z)}{(cz+d)^2}$, so

$$\phi(gz) = |f(gz)|\,\text{Im}(gz)^k = (|f(z)||(cz+d)|^{2k}) \cdot (\text{Im}(z)^k(cz+d)^{-2k}) = \phi(z).$$

But ϕ is continuous (on the fundamental domain), and as the imaginary part $y \to \infty$, we know that ϕ goes to 0 because the exponential term $e^{-2\pi y}$ dominates the polynomial term y^k. This means that ϕ is bounded, or that for all $z \in \mathbb{H}$,

$$\phi(z) \leq M \implies |f(z)| \leq My^{-k}.$$

This is helpful, because we can now extract the a_n coefficient by considering the function $\frac{f(z)}{q^{n+1}}$. If we consider $q = e^{2\pi i(x+iy)}$ for a fixed y and sending x from 0 to 1, the contour follows a circle C once counterclockwise around

$q = 0$, and thus the residue formula tells us

$$|a_n| = \left| \frac{1}{2\pi i} \int_C \frac{f(z)dq}{q^{n+1}} \right| \le \frac{1}{2\pi} \int_C \left| \frac{f(z)dq}{q^{n+1}} \right| = \int_0^1 \frac{|f(x+iy)|}{|q^n|} dx$$

(using the substitution $q = e^{2\pi i(x+iy)} \implies dq = 2\pi i q dx$). And we can now bound this with the inequality

$$|a_n| \le \int_0^1 |My^{-k} q^{-n}| dx \le My^{-k} e^{2\pi n y},$$

and taking $y = \frac{1}{n}$ gives the desired result. □

Corollary 59

If a modular form f of weight $2k$ is not a cusp form, then the coefficients have order of magnitude n^{2k-1} (that is, $|a_n| = \Theta(n^{2k-1})$).

Proof. Such modular forms of weight $2k$ can be written as $cG_k + h$, where h is a cusp form and $c \in \mathbb{C} \ne 0$. G_k's coefficients are $\Theta(n^{2k-1})$ while h's are $O(n^k)$, so the G_k coefficients dominate for large n. □

Remark 60. *Work by Pierre Deligne has shown that Theorem 58 can be improved: it has been shown that* $a_n = O(n^{k-1/2}\sigma_0(n))$*, where* $\sigma_0(n)$ *denotes the number of divisors of* n*. Since* $\sigma_0(n)$ *is subpolynomial, this tells us that for all* $\varepsilon > 0$*, we have the stronger bound* $a_n = O(n^{k-1/2+\varepsilon})$.

Serre 7.4.4 – Nikhil Reddy

Today, we'll talk a bit about the discriminant

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

where $g_2 = 60G_2$ and $g_3 = 140G_3$. To get some of the constants to work out a bit better, we can also write this in terms of E_2 and E_3: because

$$g_2 = 120\zeta(4)E_2, \quad g_3 = 280\zeta(6)E_3,$$

we can substitute these values in to find that

$$\Delta = (2\pi)^{12}(12)^{-3}(E_2^3 - E_3^2).$$

We know the q-series for E_2 and E_3, so we can directly compute the first few coefficients (which are all positive integers):

$$\Delta = (2\pi)^{12}(q - 24q^2 + 252q^3 - 1472q^4 + \cdots).$$

This power series may look a bit familiar:

Theorem 61

We have

$$\Delta = (2\pi)^{12} q \prod_{n=1}^{\infty} (1-q^n)^{24}.$$

This proof is a "bit artificial" because the natural method is to use elliptic curves.

Proof. Let $F = q\prod_{n=1}^{\infty}(1-q^n)^{24}$: it's enough to show that F is a cusp of weight 12, and then check the coefficient of the q term to figure out the scaling factor (because the space of cusp forms of weight 12 has dimension 1).

Define the two series (prime sum means we ignore $(0,0)$)

$$G_1(z) = \sum_n \sum_m{}' \frac{1}{(m+nz)^2}, \quad G(z) = \sum_m \sum_n{}' \frac{1}{(m+nz)^2}.$$

Because the double sum is not absolutely summable, the order of summation here is important. Note that G_1 is not a modular form:

Proposition 62

We have

$$G_1(z) = \frac{(2\pi)^2}{12} - 2(2\pi)^2\sum_{n=1}^{\infty}\sigma_1(n)q^n$$

and

$$G_1\left(-\frac{1}{z}\right) = z^2G_1(z) - 2\pi iz.$$

Proof. Earlier, we showed that for $k \ge 2$,

$$G_k(z) = 2\zeta(2k) + 2\frac{(2\pi i)^{2k}}{(2k-1)!}\sum_{n=1}^{\infty}\sigma_{2k-1}(n)q^n.$$

The proof basically follows the same way if you plug in $k = 1$ instead and sum in the correct order (as we do).

We'll postpone the second identity for the end. □

To show that F has weight 12, we need to show that

$$F\left(-\frac{1}{z}\right) = z^{12}F(z).$$

We'll take the logarithmic differential of both sides:

$$\frac{dF}{F} = \left(\frac{1}{q} - 24\sum_{n=1}^{\infty}\frac{nq^{n-1}}{1-q^n}\right)dq.$$

We can write this more nicely as

$$= \frac{dq}{q}\left(\frac{1}{q} - 24\sum_{n=1}^{\infty}\frac{nq^n}{1-q^n}\right)$$

and then expand as a geometric sequence to find

$$= \frac{dq}{q}\left(1 - 24\sum_{n,m=1}^{\infty}nq^{mn}\right) = \frac{dq}{q}\left(1 - 24\sum_{n=1}^{\infty}\sigma_1(n)q^n\right).$$

Using the first part of Proposition 62 yields

$$= \frac{12}{(2\pi)^2}G_1(z)\frac{dq}{q},$$

and now $\frac{dq}{q} = 2\pi i dz$ tells us that

$$\boxed{\frac{dF}{F} = \frac{12i}{2\pi}G_1(z)dz}.$$

Plugging in $-\frac{1}{z}$, we find that

$$\frac{dF\left(-\frac{1}{z}\right)}{F\left(-\frac{1}{z}\right)} = \frac{12i}{2\pi}G_i\left(-\frac{1}{z}\right)\frac{dz}{z^2},$$

and now by the second part of Proposition 62, this is

$$= \frac{12i}{2\pi}\frac{z^2G_1(z) - 2\pi iz}{z^2}dz = \left(\frac{12i}{2\pi}G_1(z) + \frac{12}{z}\right)dz.$$

But now this means that

$$\frac{dF\left(-\frac{1}{z}\right)}{F\left(-\frac{1}{z}\right)} = \frac{dF}{F} + \frac{12}{z}dz,$$

which means that $F(-\frac{1}{z}) = kz^{12}F(z)$ by reversing the logarithmic derivative. Now looking at $z = i$, we find that $F(i) = kF(i)$, so $k = 1$ (because $F(i) \neq 0$). So indeed F is a modular form of weight 12, as desired. (Checking the constant term between F and Δ just comes from looking at the q-term.) □

We can now prove the second identity

$$G_1\left(-\frac{1}{z}\right) = z^2G(z) - 2\pi iz.$$

We will need to introduce two more series

$$H_1(z) = \sum_n \sum_m{}' \frac{1}{(m-1+nz)(m+nz)}, \quad H(z) = \sum_m \sum_n{}' \frac{1}{(m-1+nz)(m+nz)}$$

where we avoid both $(1, 0)$ and $(0, 0)$. These two series can be computed directly, because the series telescopes via

$$\frac{1}{(m-1+nz)(m+nz)} = \frac{1}{m-1+nz} - \frac{1}{m+nz}.$$

So for H_1, we have the contribution from $n \neq 0$ giving us

$$H_1(z) = \sum_m \frac{1}{(m-1+nz)(m+nz)} = 0$$

and the $n = 0$ contribution just gives us 2. A more complicated calculation gives us $H(z) = 2 - \frac{2\pi i}{z}$, and now it turns out that

$$G_1\left(-\frac{1}{z}\right) = z^2G(z).$$

This is true because of how we usually write series – plugging in $\frac{1}{z}$ brings up the z to the numerator but swaps the order of summation. So now

$$\sum_{m,n}{}' \frac{1}{(m+nz)^2} - \frac{1}{(m-1+nz)(m+nz)} = \sum_{m,n}{}' \frac{1}{(m+nz)^2(m-1+nz)}$$

is absolutely convergent (order 3), so we can add it to G and G_1 to find that

$$G - H = G_1 - H_1 \implies G - G_1 = H - H_1 = \frac{2\pi i}{z},$$

which yields the result because

$$G_1\left(-\frac{1}{z}\right) = z^2\left(G_1(z) - \frac{2\pi i}{z}\right) = z^2G(z) - 2\pi iz.$$

6 February 25, 2020

Serre 7.4.5 – Michael Tang

We'll be doing some cleanup from section 2.3 (lattices and elliptic curves), and then we'll talk a bit more about the coefficients of Δ.

Definition 63

Let Γ be a lattice in the complex plane (Γ is isomorphic to $\mathbb{Z}^2$ and spans $\mathbb{C}$ over real linear combinations). Then the **Weierstrass *p*-function**

$$\wp_\Gamma(u) = \frac{1}{u^2} + \sum_{\gamma \in \Gamma} \left(\frac{1}{(u-\gamma)^2} - \frac{1}{\gamma^2} \right).$$

We're going to show that this function and its derivative satisfy the equation of an elliptic curve. To do this, we start with the Laurent expansion:

Proposition 64

The Laurent expansion of $\wp_\Gamma$ is

$$\wp_\Gamma(u) = \frac{1}{u^2} + \sum_{k=2}^{\infty} (2k-1) G_k(\Gamma) u^{2k-2},$$

where $G_k(\Gamma) = \sum'_{\gamma \in \Gamma} \frac{1}{\gamma^{2k}}$.

Sketch. There's a lot of algebra, and we can look at the 18.783 lecture notes for more details. We can expand out G_k in the proposition as a sum as well to get a double summation:

$$\frac{1}{u^2} + \sum_{k=2}^{\infty} \sum_{\gamma \in \Gamma}{}' \frac{(2k-1)u^{2k-2}}{\gamma^{2k}}$$

This converges absolutely, so we can swap the order of summation

$$= \frac{1}{u^2} + \sum_{\gamma \in \Gamma}{}' \sum_{k=2}^{\infty} (2k-1) \frac{u^{2k-2}}{\gamma^{2k}}.$$

For each k, this is an "arithmetico-geometric series" which we can evaluate directly. But to get the exact form of the Weierstrass p function, we need to do a trick: we actually add in the "odd terms" with γ^{2k+1} back in (they cancel out because γ^n cancels with $(-\gamma)^n$), and that will give us what we want. □

Proposition 65

Let $x = \wp_\Gamma(u)$ and $y = \wp'_\Gamma(u)$. Then

$$y^2 = 4x^3 - g_2 x - g_3,$$

where $g_2 = 60G_2(\Gamma)$ and $g_3 = 140G_3(\Gamma)$.

Sketch. Define the function

$$F(u) = y^2 - (4x^3 - g_2 x - g_3):$$

write out the Laurent expansion and bash, and we'll see that the negative-exponent terms cancel out, and also $F(0) = 0$. This means that F is a holomorphic function at $z = 0$.

Now $\wp_\Gamma$ and $\wp'_\Gamma$ are both periodic with respect to the lattice, so F is also periodic. And this means that F is holomorphic at all points of the lattice Γ: a uniform convergence argument shows that $F, \wp_\Gamma, \wp'_\Gamma$ are all holomorphic at all points not on the lattice, so F is also holomorphic everywhere.

Now F is periodic with no poles and entire, so it is bounded by the supremum on one (compact) fundamental domain, so it must be constant by Liouville, and thus $F = 0$ everywhere. □

This is the **Weierstrass form** for elliptic curves – we can prove that it is nonsingular, so there is something about an isomorphism between curves and lattices (though we don't have the background).

Next, we'll go back to the modular form

$$\Delta = g_2^3 - 27g_3^2.$$

Recall a few important properties here: Δ has weight 12, and it's a cusp form (this explains the coefficient 27). Also, it has the q-expansion

$$\Delta = (2\pi)^{12} q \prod_{n=1}^{\infty} (1 - q^n)^{24}.$$

Definition 66

The **Ramanujan τ function** is defined such that

$$q \prod_{n=1}^{\infty} (1 - q^n)^{24} = \sum \tau(n) q^n.$$

We can check some small values: $\tau(1) = 1, \tau(2) = -24, \tau(3) = 252, \tau(4) = -1472$ grow pretty quickly, but we can give a bound on the coefficients. Recall that a cusp form of weight k has q-coefficients that are $O(n^k)$ (and in fact $O(n^{k-1/2+\varepsilon})$. So in this case, $\tau(n) = O(n^6)$ and actually $O(n^{11/2+\varepsilon})$.

The τ function has nice properties, but we'll prove them later on:

Proposition 67

We have the following results:

- τ is multiplicative: for any m, n relatively prime, $\tau(mn) = \tau(m)\tau(n)$. (This means we only need the values at prime powers.)
- We have the second-order recursion (for prime p and $n > 1$)
$$\tau(p^{n+1}) = \tau(p)\tau(p^n) - p^{11}\tau(p^{n-1})$$
- We can write the Dirichlet series $L(\tau, s) = \sum_{n=1}^{\infty} \frac{\tau(n)}{n^s}$ as a product
$$L(\tau, s) = \prod_{p \text{prime}} \frac{1}{1 - \tau(p)p^{-s} + p^{11-2s}}.$$

It's conjectured (unproven) that $\tau(n) \neq 0$ for all n, but this has been checked up to about 8×10^{23}.

Serre 7.5.1 – Natalie Stewart

We'll start by redefining a few familiar concepts:

Definition 68

Let E be a set. Then X_E denotes the **free abelian group** on E:

$$X_E = \left\{ \sum_{x \in E} c_x x | c_x \in \mathbb{Z}, \text{all but finitely many zero} \right\}.$$

Definition 69

A **correspondence** on a set E is an endomorphism $T : X_E \to X_E$ – an equivalent way to think about this is that we can write

$$T(x) = \sum_{y \in E} n_T(x,y) y,$$

which is equivalent to a set function $n_T : E \to \mathbb{N}^E$ with finite support.

Definition 70

Let f be a function from a set E to a group G. Define $Tf : E \to G$ as the composition

$$E \hookrightarrow X_E \xrightarrow{T} X_E \xrightarrow{f} G.$$

Note that the endomorphisms on (say) an abelian group can be added pointwise and multiplied via composition, so they have a structure like a ring. We'll be defining a bunch of correspondences which basically form a subring.

Definition 71

Let $\mathcal{R}$ be the set of lattices on $\mathbb{C}$. Define the correspondence $T(n)$ by

$$T(n)\Gamma = \sum_{\substack{\Gamma' \in R \\ (\Gamma : \Gamma') = n}} \Gamma'.$$

for all $n > 1$. For all $\lambda \in \mathbb{C}^\times$, define the map from $X_\mathcal{R}$ to $X_\mathcal{R}$ via

$$R_\lambda \Gamma = 1 \cdot (\lambda \Gamma),$$

where $(\lambda\Gamma)$ denotes the homothetic rescaling of Γ.

Note that if $\Gamma' \in \Gamma$ (is a sublattice) has index n, then Γ' will contain $n\Gamma$. This gives us a chain of inclusions, and

$$\{\Gamma' | (\Gamma : \Gamma') = n\} \cong \{G \in \Gamma / n\Gamma | |G| = n\}.$$

We'll be using this a lot in the upcoming proof. This reduces the job of counting the subgroups of a certain order for the group $\Gamma/n\Gamma$. We'll use the fact that if n is a prime, then this number is actually $n + 1$.

Proposition 72

For all $\lambda, \mu \in \mathbb{C}^\times$ and $n, m \in \mathbb{Z}_{>1}$, we have the following:

1. Homothety operators multiply: $R_\lambda R_\mu = R_{\lambda\mu}$.
2. $R_\lambda T(n) = T(n) R_\lambda$ (we have commutativity).
3. $T(n)T(m) = T(nm)$ if n, m are relatively prime.
4. $T(p^n)T(p) = T(p^{n+1}) = pT(p^{n-1})R_p$ if p is a prime.

Proof. (1) and (2) are clear (arguments about scaling).

For (3), note that if $(\Gamma : \Gamma'') = nm$, then the coefficient of Γ'' in $T(m)T(n)\Gamma$ comes from counting the lattices Γ' where $\Gamma'' \subset \Gamma' \subset \Gamma$ and $(\Gamma : \Gamma') = n$, and by an isomorphism theorem this is just the number of subgroups G of Γ/Γ'' with $|G| = m$. Since m, n are relatively prime, it's sufficient to check that the coefficient of Γ'' in $T(n)T(m)\Gamma$ is 1, which means there is a unique subgroup of Γ/Γ'' of order m (this is true).

For (4), note that a sublattice is represented only if its index is p^{n+1}, so we can just check that the coefficients are the same there. Consider Γ'' such that $(\Gamma : \Gamma'') = p^{n+1}$: from an earlier argument, Γ' with index p must contain $p\Gamma$.

Suppose that the coefficient of Γ'' in $T(p^n)T(p)$ is a, and the coefficient in $T(p^{n-1})R_p$ is c. we wish to show that $a = 1 + pc$. We break into cases:

(case i) If Γ'' is not contained in $p\Gamma$, then $c = 0$, and we want to show that $a = 1$. Note that $|\Gamma/p\Gamma| = p^2$, so the order of the image of the projection from Γ into $\Gamma/p\Gamma$ is p. Also, because Γ'' is not contained in $p\Gamma$, it doesn't intersect at all, and the order of the image of Γ'' is also p. Since the image of Γ'' is contained in the image of Γ', we can use the uniqueness part of the isomorphism theorem to show that we have a bijective map between subgroups – there exists a unique Γ'.

(case ii) If Γ'' is contained in $p\Gamma$, which is contained in Γ', we can check that $c = 1$ and $a = p + 1$ (we did this above). □

Corollary 73

Inductively, we can show that the prime powers $T(p^n)$ are polynomials in the underlying prime $T(p)$ and the homothety operator R_p. Also, the ring generated by $T(p)$ and R_λ for p prime and λ contains $T(n)$ for all $n > 1$ and its commutative.

One important idea is that these correspondence act on our lattice functions: recall that a lattice function of weight $2k$ from $\mathcal{R}$ to $\mathbb{C}$ satisfies

$$F(\lambda\Gamma) = \lambda^{-2k} F(\Gamma).$$

Using the homothety operator, we can write that

$$F(R_\lambda \Gamma) = \lambda^{-2k} F(\Gamma),$$

so $T(n)F$ is also a lattice function of weight $2k$. And this gives us an extra characterization similar to the above proposition:

$$T(m)T(n)F = T(mn)F, \quad T(p^n)T(p)F = T(p^{n+1})F + p^{1-2k}T(p^{n-1})F.$$

Serre 7.5.2 – Kaarel Haenni

We'll be talking about sublattices of a particular lattice. Throughout this, we'll be fixing a lattice Γ with basis ω_1, ω_2: define $\Gamma(n)$ to be the set of sublattices of Γ of index n.

Definition 74

Let S_n be the set of integer matrices $\begin{bmatrix} a & b \\ 0 & d \end{bmatrix}$ with $ad = n$, $a \geq 1$, and $0 \leq b < d$.

Definition 75

For any $\sigma = \begin{bmatrix} a & b \\ 0 & d \end{bmatrix}$, define the sublattice Γ_σ be the sublattice with basis $\omega_1' = a\omega_1 + b\omega_2$ and $\omega_2' = d\omega_2$.

Theorem 76

The map $\sigma \to \Gamma_\sigma$ is a bijection from S_n to $\Gamma(n)$.

Proof. First of all, we check that the index is indeed n: this is equal to the ratio of areas of a fundamental cell of Γ_σ to Γ, which is indeed $\det \sigma = ad = n$. (These two equalities follow by looking at the "cross product" of two vectors in the complex plane and considering the number of lattice points in one fundamental domain.)

To show that we have a bijection, we construct an inverse map sending a sublattice Γ' to $\sigma(\Gamma')$. Define the two groups

$$Y_1 = \Gamma/(\Gamma' + \mathbb{Z}\omega_2), \quad Y_2 = \mathbb{Z}\omega_2/(\Gamma' \cap \mathbb{Z}\omega_2).$$

Both of these are cyclic groups, generated by ω_1, ω_2 in the quotient maps. Let a and d be the orders of these cyclic groups ω_1, ω_2. Now defining $\omega_2' = d\omega_2$, we know that $\omega_2' \in \Gamma'$. Also, by definition of a, we know that there is some unique b such that $a\omega_1 + b\omega_2 \in \Gamma'$ as long as $0 \leq b < d$. Then let $\omega_1' = a\omega_1 + b\omega_2$. Now we have a, b, d and we can just define

$$\sigma(\Gamma') = \begin{bmatrix} a & b \\ 0 & d \end{bmatrix}.$$

Showing that these maps are inverses is routine, so we'll skip it here. □

Example 77

Let p be a prime. Then the sublattices of Γ of index p can be easily characterized: S_p consists of the matrices $\begin{bmatrix} p & 0 \\ 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & b \\ 0 & p \end{bmatrix}$ for all $0 \leq b < p$.

(This shows that the statement from Natalie's lecture that there are $p+1$ sublattices of index p.) These correspond to the sublattices $\omega_1' = p\omega_1, \omega_2' = \omega_2$ (scaling one variable by p) and $\omega_1' = \omega_1 + b\omega_2, \omega_2' = p\omega_2$ (scaling the other by p and doing a shear).

7 February 27, 2020

Serre 7.5.3, 7.5.4 – Michelle Xu

We recently introduced the operator $T(n)$, which acts on lattices as follows:

$$T(n)\Gamma = \sum_{(\Gamma:\Gamma')=n} \Gamma'.$$

Today, we'll talk about this in the context of modular functions.

Recall that we have a map between **lattice functions** F of weight $2k$ and **weakly modular functions** f of weight $2k$, with the correspondence

$$\omega_2^{2k} F(\Gamma(\omega_1, \omega_2)) = f\left(\frac{\omega_1}{\omega_2}\right).$$

Definition 78

$T(n)$ acts on functions f with the relation

$$T(n)f(z) = n^{2k-1} T(n) F(z, 1).$$

(The n^{2k-1} is there for aesthetic reasons). Remember that the set of sublattices of index n for a lattice $\Gamma(\omega_1, \omega_2)$ can be represented with the set of bases

$$\begin{bmatrix} \omega_1' \\ \omega_2' \end{bmatrix} = \begin{bmatrix} a & b \\ 0 & d \end{bmatrix} \begin{bmatrix} \omega_1 \\ omega_2 \end{bmatrix},$$

where $ad = n, 0 \le b < d, a \ge 1$. This means we can also rewrite our equation above as

$$\boxed{T(n)f(z)n^{2k-1} \sum_{(\Gamma:\Gamma')=n} F(\Gamma') = n^{2k-1} \sum_{a,d,b} d^{-2k} f\left(\frac{az+b}{d}\right)}$$

(expressing in terms of the bases).

Proposition 79

Let f be a weakly modular function of weight $2k$. Then the function $T(n)f$ is also weakly modular of weight $2k$, and it is holomorphic if f is.

Proof. Remember that the action on $SL_2(\mathbb{Z})$ on our lattice function keeps it invariant, so

$$T(n)f(z) = n^{2k-1} T(n) F(z, 1) = n^{2k-1} T(n) F(az + b, cz + d),$$

and converting this back to our weakly modular functions gives us

$$= (cz + d)^{-2k} T(n) f\left(\frac{az+b}{cz+d}\right),$$

so this is indeed of weight $2k$. And now we have a finite number of terms in the boxed equation above, each of which is meromorphic (because f is meromorphic). So $T(n)f$ is also meromorphic, and it is holomorphic if f is (for the same reason). □

Proposition 80

Our operator $T(n)$ satisfies the following equations:

- $T(n)T(m)f = T(mn)f$ for m, n relatively prime.
- $T(p)T(p^n)f = T(p^{n+1})f + p^{2k-1}T(p^{n-1})f$ for p prime and $n \ge 1$.

Proof. Recall that we proved a very similar set of relations for lattice functions – we just plug this in along with our definition of an operator.

The main point is that the p^{2k-1} factor is different; this mostly comes from the n^{2k-1} definition of our operator on modular functions. □

Proposition 81

Let f be a modular function of weight $2k$ (so it is also meromorphic at infinity) with Laurent expansion $f(z) = \sum_{m\in\mathbb{Z}} c(m)q^m$. Then $T(n)f$ is also a modular function, we can write

$$T(n)f = \sum_{m\in\mathbb{Z}} \gamma(m)q^m,$$

where $\gamma(m) = \sum_{a\ge1, a|(n,m)} a^{2k-1}c\left(\frac{mn}{a^2}\right)$.

Proof. We expand f with its Laurent series to write

$$T(n)f = n^{2k-1}\sum_{\substack{ad=n\\ a\ge1\\ 0\le b<d}} d^{-2k}\sum_{m\in\mathbb{Z}} c(m)e^{2\pi im(az+b)/d}.$$

(here we have expanded $q = e^{2\pi iz}$).

Fix d and m, and consider the sum $\sum_{0\le b<d} e^{2\pi imb/d}$. If $d|m$, then the exponent is a multiple of $2\pi i$, so this sum is just d. Otherwise, we have a sum over the roots of unity, which is 0. To use this, we define $\frac{m}{d} = m'$, and then we can write

$$T(n)f = n^{2k-1}\sum_{\substack{ad=n\\ a\ge1\\ m'\in\mathbb{Z}}} d^{-2k+1}c(m'd)q^{am'}.$$

To simplify further, let $\mu = am'$, and now we just have a single exponent for q:

$$T(n)f(z) = \sum_{\mu\in\mathbb{Z}} q^{\mu}\sum_{a|(n,\mu), a\ge1}\left(\frac{n}{d}\right)^{2k-1} c\left(\frac{\mu d}{a}\right).$$

And now this is basically what we want by relabeling indices – we just need to show $T(n)f$ is meromorphic. Because f is meromorphic, its Laurent expansion stops at some N: $c(m) = 0$ for all $m \le N$. This means that $c\left(\frac{\mu d}{a}\right) = 0$ for all $\mu \le -nN$ (because the largest $\frac{d}{a}$ can be is n). This means that f is meromorphic at infinity, and because it is weakly modular (from earlier arguments), $T(n)f$ is indeed a modular function, as desired. □

So $T(n)$ brings modular functions to modular functions, which is nice.

Corollary 82

Using the same notation as above, we have $\gamma(0) = \sigma_{2k-1}(n)c(0)$, where $\sigma_{2k-1}(n)$ is the sum of the $(2k-1)$th divisors of n, and $\gamma(1) = c(n)$. Also, if $n = p$ is prime, then we have

$$\gamma(m) = \begin{cases} c(pm) & m \neq 0 \pmod p \\ c(pm) + p^{2k-1}c\left(\frac{m}{p}\right) & m = 0 \pmod p. \end{cases}$$

(These can be verified by directly plugging in $m = 0, 1$ and $n = p$ and using divisibility properties.)

Corollary 83

If f is a modular form (resp. cusp form), then $T(n)f$ is modular (resp. cusp) as well.

Proof. The argument follows analogously as above, replacing "meromorphic at infinity" with "holomorphic at infinity." Cusp form follows from the fact that if $c(0) = 0$, then $\gamma(0) = 0$ from the above corollary. □

With this, we'll study something more specific about our $T(n)$: eigenvalues and eigenfunctions. We'll assume that $f = \sum_{n=0}^{\infty} c(n)q^n$ is a modular form of weight $2k$ for some $k > 0$.

Definition 84

f is an **eigenfunction** for all $T(n)$ if we have $\lambda(n) \in \mathbb{C}$ such that

$$T(n)f = \lambda(n)f.$$

(Then the $\lambda(n)$ are the set of **eigenvalues** for $T(n)$.)

Theorem 85

Let $f = \sum_{n=0}^{\infty} c(n)q^n$ be an eigenfunction. Then the coefficient of $c(1)$ is nonzero, and if we normalize $c(1) = 1$, we have $c(n) = \lambda(n)$.

Proof. We know that the coefficient of q for $T(n)f$, which is $\gamma(1)$, is equal to $c(n)$. But by the definition of the eigenfunction, $T(n)f$ should have q-coefficient $\lambda(n)c(1)$, so $c(n) = \lambda(n)c(1)$. Then $c(1)$ can't be zero, because $c(n) = 0$ for all $n \geq 1$ (and that means f is a constant, and the only modular form for $k > 0$ is trivial). Thus $c(1) \neq 0$, and the second claim follows easily by setting $c(1) = 1$. □

Corollary 86

If two modular forms of weight $2k$ are eigenfunctions of $T(n)$ with the same eigenvalues $\lambda(n)$, and they are both normalized (to $c(1) = 1$), then they coincide.

(This is because the Laurent expansion is completely defined by the eigenvalues.)

This allows us to say something more specific about our eigenfunctions f:

Corollary 87

Suppose that f is an eigenfunction of all $T(n)$, and it is normalized to $c(1) = 1$. Then

$$c(m)c(n) = c(mn) \quad \forall (m, n) = 1,$$

$$c(p)c(p^n) = c(p^{n+1}) + p^{2k-1}c(p^{n-1}).$$

(We plug in the properties of our operator $T(n)$ into the eigenfunction statement.)

Definition 88

Let f be an eigenfunction. Then

$$\Phi_f(s) = \sum_{n=1}^{\infty} \frac{c(n)}{n^s}$$

is the Dirichlet series defined by the coefficients of an eigenfunction f.

This series converges absolutely for all $\mathrm{Re}\, s > 2k$, because we proved that modular forms' coefficients have order of magnitude $c(n) = O(n^{2k-1})$.

Corollary 89

Let P be the set of prime numbers. Then

$$\Phi_f(s) = \prod_{p \in P} \frac{1}{1 - c(p)p^{-s} + p^{2k-1-2s}}.$$

Proof. First, we want to show that we can write our Dirichlet series as

$$\Phi_f(s) = \prod_{p \in P} \left(\sum_{n=0}^{\infty} c(p^n)p^{-ns} \right).$$

(This proof comes from Serre chapter 6.) Let S be a finite set of prime numbers, and let $N(S)$ be the set of integers with prime factors only in S. We know that coefficients are multiplicative, so we can rewrite

$$\sum_{n \in N(s)} \frac{c(n)}{n^s} = \prod_{p \in S} \left(\sum_{m=0}^{\infty} c(p^m)p^{-ms} \right).$$

Now as S approaches the set of all primes, our infinite product will converge to the result above. From this, we just need to show the inner term works out: in other words, for all primes p, we want to show (defining $Q = p^{-s}$)

$$\sum_{n=0}^{\infty} c(p^n)Q^n = \frac{1}{1 - c(p)Q + p^{2k-1}Q^2}.$$

To do this, we consider the series

$$\psi(n) = (\sum_{n=0}^{\infty} c(p^n)Q^n)(1 - c(p)Q + p^{2k-1}Q^2);$$

we wish to show that this is equal to 1. Now the Q-coefficient is equal to $c(p) - c(p) = 0$, and Q^{n+1} for all $n \geq 1$ has coefficient (expanding out the relevant terms)

$$c(p^{n+1}) - c(p)c(p^n) + p^{2k-1}c(p^{n-1}),$$

which is zero because of the recursive equation for $c(p^n)$. And this means that we only have a constant term, and $\psi(n) = c(1) = 1$ by normalization, completing the proof. □

Hecke also proved a few interesting facts about $\Phi_s(f)$ – this will have a meromorphic continuation over the plane, just like the Riemann zeta function.

Serre 7.5.5, 7.5.6 – Shreyas Balaji

Throughout this section, "normalized eigenfunction" means $c(1) = 1$. We'll do a few examples of eigenfunctions for the Hecke operators $f(n)$ first:

Proposition 90

The Eisenstein series G_k is an eigenfunction of $T(n)$ with eigenvalue $\lambda = \sigma_{2k-1}(n)$. Specifically, the normalized eigenfunction is

$$f = (-1)^k \frac{B_k}{4k} E_k = (-1)^k \frac{B_k}{4k} + \sum_{n=1}^{\infty} \sigma_{2k-1}(n) q^n,$$

where B_k are the Bernoulli numbers.

Proof. Let $G_k(\Gamma)$ be the function from the set of lattices $\mathcal{R}$ to the complex numbers, where

$$G_k(\Gamma) = \sum_{\gamma \in \Gamma}{}' \frac{1}{\gamma^{2k}}.$$

We can just show that Eisenstein series are eigenfunctions for all $T(p)$ for prime p (and then use the relations to build up to all $T(n)$). Then for any prime p,

$$T(p) G_k(\Gamma) = \sum_{(\Gamma:\Gamma')=p} \sum_{\gamma \in \Gamma}{}' \frac{1}{\gamma^{2k}}.$$

Consider a particular $\gamma \in \Gamma$: then there are two cases. **(1)** We have $\gamma \in p\Gamma$ (that is, the lattice vectors scaled by p). There are $p + 1$ sublattices that have index p in Γ, and if it's inside $p\Gamma$, it's inside all of the sublattices as well. **(2)** γ is contained in exactly one sublattice. So we can split these up to find that

$$T(p) G_k(\Gamma) = G_k(\Gamma) + p G_k(p\Gamma)$$

(the first term from every point being counted once, and the second coming from points that are counted $p+1$ times). And note that $G_k(p\Gamma) = \gamma^{-2k} G_k(\Gamma)$, so

$$T(p) G_k(\Gamma) = (1 + p^{1-2k}) G_k(\Gamma),$$

which shows that $G_k(\Gamma)$, viewed as a function on lattices, is an eigenfunction for the operator $T(p)$. And now this means the modular form G_k associated with the function $G_k(\Gamma)$ is an eigenfunction under $T(p)$ with eigenvalue $p^{2k-1}(1 + p^{1-2k}) = \sigma_{2k-1}(p)$, as desired. (We pick up a p^{2k-1} factor when we convert between modular functions and modular forms.) □

Proposition 91

Let f be defined as above. Then the Dirichlet series $\Phi_f(s) = \zeta(s)\zeta(s - 2k + 1)$.

Proof. We can write out the coefficients

$$\Phi_f(s) = \sum_{n=1}^{\infty} \frac{\sigma_{2k-1}}{n^s} = \sum_{a,d \geq 1} \frac{a^{2k-1}}{a^s d^s}.$$

We can now separate this into two different sums:

$$= \left(\sum_{d \geq 1} \frac{1}{d^s} \right) \left(\sum_{a \geq 1} \frac{1}{a^{s+1-2k}} \right),$$

which is exactly what we want. □

Proposition 92

The Δ function is an eigenfunction of $T(n)$ with normalized eigenfunction

$$(2\pi)^{-12}\Delta = q \prod_{n=1}^{\infty} (1-q^n)^{24} = \sum_{n=1}^{\infty} \tau(n) q^n.$$

Proof. Remember that the space of cusp forms of weight 12 is of dimension 1, and this space is stable under action of $T(n)$. Thus $T(n)\Delta$ must be some constant times Δ, and we can check the coefficients for normalization. □

Corollary 93

The τ function satisfies $\tau(nm) = \tau(n)\tau(m)$ for all relatively prime n, m, and $\tau(p)\tau(p^n) = \tau(p^{n+1}) + p^{11}\tau(p^{n-1})$ for all prime p and $n \geq 1$.

(This follows from directly applying results from above about the coefficients $c(n)$ of eigenfunctions.)

We'll now move on to complements: we will be stating but not proving a lot of results.

Definition 94

Let f, g be cusp forms of weight $2k$. Then define a measure μ via

$$d\mu(f, g) = f(z)\overline{g(z)} y^{2k} dx dy / y^2,$$

where x, y are the real and imaginary part of z.

Note that μ is invariant under action by G (because of the y^{2k} term) and it is bounded over H/G (because f, g are bounded and f and g fall off exponentially).

Definition 95

The **Petersson scalar product** on $\mathcal{M}_k^0$ is defined as

$$\langle f, g \rangle = \int_{H/G} d\mu(f, g) = \int_D f(z)\overline{g(z)} y^{2k-2} dx dy$$

where D is the fundamental domain.

Fact 96

The operators $T(n)$ turn out to be Hermitian under the Petersson scalar product: we have

$$\langle T(n)f, g\rangle = \langle f, T(n)g\rangle.$$

Since $T(n)$ commute with each other, the operators are simultaneously diagonalizable. Furthermore, the image of f is contained within the span of the $T(n)$ eigenvectors (the operators are **complete**), so **there exists an orthogonal basis of $\mathcal{M}_k^0$ made of eigenvectors of $T(n)$ with real eigenvalues** (because the operators are Hermitian).

We'll move on to something completely unrelated:

Definition 97

Let $M_k(\mathbb{Z})$ be the set of weight $2k$ modular forms which can be written as $f = \sum_{n=0}^{\infty} c_n q^n$ for c_n **integers**.

It turns out that there's a specific $\mathbb{Z}$-basis of $M_k(\mathbb{Z})$ (it must exist because the space is stable under action of $T(n)$), which extends to a $\mathbb{C}$-basis of M_k. Specifically, we pick

$$\{E_2^{\alpha} F^{\beta} : \alpha + 3\beta = k/2\}$$

for k even, and

$$\{E_3 E_2^{\alpha} F^{\beta} : \alpha + 3\beta = (k-3)/2\}$$

for k odd. The nice thing is that this $\mathbb{Z}$-basis extends to a $\mathbb{C}$-basis, so the coefficients of the characteristic polynomial of $T(n)$ **must be integers** (because they come from a $\mathbb{Z}$-basis).

Fact 98 (Recently proven Petersson conjecture)

Let $f = \sum_{n \ge 1} c(n) q^n$ be a cusp form of weight $2k$ which is also a normalized eigenfunction for $T(n)$. Define

$$\Phi_{f,p}(T) = 1 - c(p)T + p^{2k-1}T^2 :$$

then we can factor this polynomial as $(1 - \alpha_p T)(1 - \alpha_p' T)$, where $\alpha_p + \alpha_p' = c(p), \alpha_p \alpha_p' = p^{2k-1}$. Then α_p, α_p' are complex conjugates.

8 March 3, 2020

Serre 7.6.1-7.6.4 – Vanshika Jain

We'll start with a few definitions that we'll need throughout the rest of this lecture:

Definition 99

An **invariant measure** μ is a measure preserved under translation and rotation.

For example, the normal "product measure" $dx_1 dx_2 \cdots dx_n$ is an invariant measure.

Definition 100

The **dual** of a vector space V is the space of linear functionals (linear maps from V to the scalar field) on V, with structure of pointwise addition and scalar multiplication.

Definition 101

A **rapidly decreasing** (smooth) function is a function $f(x)$ such that $f, f', f'', \cdots$ exist everywhere and decay faster than any negative power of x.

($f(x) = e^{-x}$ is such an example.) Note that the space of rapidly decreasing, smooth functions have the property that the Fourier transform is an **automorphism** of the space.

Let V be a real vector space of dimension n, equipped with an invariant measure μ. Let V' be the dual of V, and let f be a rapidly decreasing smooth function of V. Then the **Fourier transform** of f is defined to be

$$\hat{f}(y) = \int_V e^{-2i\pi\langle x,y\rangle} f(x)\mu(x).$$

If we let Γ be a lattice in V, and let Γ' be the **dual lattice** in V' (that is, the set of $y \in V'$ such that $\langle x, y\rangle = y(x)$ is an integer). Since V is a finite-dimensional vector space, there exists an isomorphism between V and V' given by

$$v \mapsto \langle v, \cdot\rangle$$

for an inner product $\langle \cdot, \cdot\rangle$.

Example 102

If we let $V = \mathbb{R}^n, \Gamma = \mathbb{Z}^n$, the dual vector space is isomorphic to $\mathbb{R}^n$, and the lattice Γ' is just $\mathbb{Z}^n$ if we work with the usual dot product.

Proposition 103

Let $v = \mu(V/\Gamma)$ (the measure of one fundamental region of Γ), and let f be a rapidly decreasing smooth function. Then

$$\sum_{x\in\Gamma} f(x) = \frac{1}{v}\sum_{y\in\Gamma'} \hat{f}(y).$$

Proof. Replace μ with a scalar multiple so that the measure of the fundamental domain becomes 1. If we take a basis $e_1, \cdots, e_n$ of Γ, we can then identify V with $\mathbb{R}^n$, Γ with $\mathbb{Z}^n$, and take μ to be the ordinary product measure. Then the formula we're trying to show just reduces to the classical Poisson summation formula

$$\sum_{x\in\mathbb{Z}^n} f(x) = \sum_{x\in\mathbb{Z}^n} \hat{f}(x),$$

which is a property of the Fourier transform. □

We'll now apply this to quadratic forms: suppose that V is equipped with a symmetric bilinear form $\langle x, y\rangle$ that is positive and nondegenerate. As before, this defines an isomorphism between V and V', so the dual lattice Γ' can be thought of as a lattice in V instead of V': a point $y \in \Gamma'$ if and only if $\langle x, y\rangle$ is an integer for all $x \in \Gamma$.

Definition 104

Associate a function to each lattice Γ via

$$\Theta_\Gamma(t) = \sum_{x \in \Gamma} e^{-\pi t \langle x, x \rangle}.$$

Proposition 105

We have the relation (for all $t \in \mathbb{R}_{\ge 0}$ and lattices Γ)

$$\Theta_\Gamma(t) = t^{-n/2} v^{-1} \Theta_{\Gamma'}(t^{-1}).$$

Proof. Let $f(x) = e^{-\pi \langle x, x \rangle}$: this is rapidly decreasing and smooth, so if we pick an orthonormal basis to identify V with $\mathbb{R}^n$ and μ with the product measure, we have

$$f = e^{-\pi(x_1^2 + x_2^2 + \cdots + x_n^2)}.$$

Since the Fourier transform $e^{-\pi x^2}$ is itself, we can use the previous result: note that $t^{1/2}\Gamma$ has dual lattice $t^{-1/2}\Gamma'$, and the volume of our lattice is $t^{n/2}v$. Thus

$$\sum_{x \in t^{1/2}\Gamma} f(x) = \sum_{x \in t^{1/2}\Gamma} e^{-\pi(x_1^2 + \cdots + x_n^2)} = \sum_{x \in \Gamma} e^{-\pi t (x_1^2 + \cdots + x_n^2)} = \Theta_\Gamma(t).$$

But we can use the Poisson summation formula and do the same trick with $t^{-1/2}\Gamma'$ to find that this is also equal to $t^{-n/2} v^{-1} \Theta_{\Gamma''}(t^{-1})$, as desired. □

Everything with Θ here can be represented using a matrix: if we let $e_1, \cdots, e_n$ be a basis for our lattice Γ, we can define $a_{ij} = \langle e_i, e_j \rangle$, which gives us a positive, symmetric, nondegenerate matrix $A = (A_{ij})$. Thus, for any $x \in V$, we can write

$$\langle x, x \rangle = \sum_{i,j} a_{ij} x_i x_j,$$

so we can write our function

$$\Theta_\Gamma(t) = \sum_{x_i \in \mathbb{Z}} e^{-\pi t \sum_{i,j} a_{ij} x_i x_j}.$$

Defining the determinant using the wedge product, we can see that the volume v of Γ is the square root of $\det A$. Furthermore, if $B = (B_{ij})$ is the inverse matrix, we have a dual basis

$$e_i' = \sum_j b_{ij} e_j.$$

Then the e_i' form a basis of Γ' – we have $\langle e_i', e_j' \rangle = B_{ij}$. In particular, this tells us that if $v' = \mu(V/\Gamma')$, we have $vv' = 1$.

For the upcoming section, we'll be dealing with some special cases – pairs (V, Γ) with a few nice properties. First of all, we want Γ' to be equal to Γ, which is equivalent to saying that $\langle x, y \rangle$ is an integer for all $x, y \in \Gamma$. This also implies that Γ's matrix A has integer coefficients, and A has determinant 1.

We'll also assume that $\langle x, x \rangle$ is always even – this means that Γ is of type II. We'll see why this is useful soon.

Serre 7.6.5-7.6.7 – Andrew Lin

In these final sections of the book, we'll discuss a particular kind of modular form with particularly nice properties. Much of the background work has already been done: throughout this lecture, we'll assume that we have a lattice Γ on a vector space V of dimension n, such that the fundamental domain has volume 1 and the diagonal entries of the matrix A_{ij} are even (we have a **type II lattice**). Since any $x \in \Gamma$ can be written as a $\mathbb{Z}$-combination of basis elements $\sum_i x_i e_i$, we know that

$$\langle x, x\rangle = \sum_i x_i^2 a_{ii} + 2\sum_{i<j} x_i x_j a_{ij}$$

is **always even** for a type II lattice.

Definition 106

For a lattice Γ, let $r_\Gamma(m)$ be the number of elements $x \in \Gamma$ such that $\langle x, x\rangle = 2m$.

We know that our bilinear form is positive and nondegenerate, and thus it is essentially a rescaling of the dot product (because we can diagonalize the matrix A that defines it). Thus, $r_\Gamma(m)$ is bounded by a constant (depending on Γ) times the number of points in $\mathbb{Z}^n$ of distance at most $\sqrt{2m}$ from the origin. Such points are contained in the n-dimensional box of side length $2\sqrt{2}m$: thus $r_\Gamma(m) = O(m^{n/2})$ for any Γ, so $r_\Gamma(m)$ is polynomial in m. This means the infinite sum

$$\sum_{m\geq 0} r_\Gamma(m) q^m$$

is convergent for all $|q| < 1$. (It's good to make sure it's clear why the constant term $r_\Gamma(0) = 1$ for any lattice Γ.)

Here's where we'll connect this back to our modular forms: notice that what we've written down is the Fourier expansion for a modular form on the upper half-plane, where $q = e^{2\pi i z}$.

Definition 107

The **theta function** associated to a lattice Γ is

$$\theta_\Gamma(z) = \sum_{m\geq 0} r_\Gamma(m) e^{2\pi i m z} = \sum_{x\in\Gamma} e^{\pi i z\langle x,x\rangle}.$$

Because this sum is absolutely convergent and we have no poles, θ_Γ is holomorphic on the upper half-plane.

Theorem 108

Suppose Γ satisfies the above assumptions. Then

(a) n is a multiple of 8,

(b) θ_Γ is a modular form of weight $\frac{n}{2}$.

Proof. The proof of the first point is outside the scope of this lecture [1], but we'll do a somewhat circular proof of (a) once we prove (b).

We know that θ_Γ is holomorphic, and it satisfies the relation under T because we've written it as a Fourier series.

[1] See chapter 5 of Serre. The main idea is that the signature of our bilinear form can be essentially related to a canonical element of the dual lattice of $(\Gamma/2\Gamma)^n$, if we look at images of the dot product mod 8.

Thus, it suffices to check the relation under S: to do this, we will instead check that

$$\boxed{\theta_\Gamma\left(-\frac{1}{z}\right) = (iz)^{n/2}\theta_\Gamma(z)}.$$

(The extra $i^{n/2}$ factor cancels out if n is a multiple of 8.) To show this result, note that both sides are holomorphic on the entire half-plane, so it suffices to show that they are equal on a set containing a limit point. Specifically, plugging in $z = it$, note that the left hand side evaluates to (using $\theta_\Gamma(r) = \sum_{x\in\Gamma} e^{\pi iz\langle x,x\rangle}$)

$$\theta_\Gamma\left(-\frac{1}{z}\right) = \sum_{x\in\Gamma} e^{-\pi/t\langle x,x\rangle} = \Theta\left(\frac{1}{t}\right),$$

while the right hand side is

$$(it)^{n/2}\theta_\Gamma(z) = t^{n/2}\sum_{x\in\Gamma} e^{-\pi t\langle x,x\rangle} = t^{n/2}\Theta(t).$$

(Here, we use the definition of $\Theta(z) = \sum e^{-\pi t\langle x,x\rangle}$ from a previous section.) These expressions are indeed equal (see Serre 7.6.2) as long as the volume v associated to our lattice is 1, which is one of our assumptions. Thus we've shown the boxed result, and thus the S-relation holds: θ_Γ is indeed a modular form.

With this, we can verify that n must be a multiple of 8 (in a circular manner). Suppose not (for the sake of contradiction); then we can assume $n \equiv 4 \pmod 8$ by replacing Γ with $\Gamma \oplus \Gamma$ (either once or twice) and still have a (theoretically) valid lattice. The boxed equation then becomes

$$\theta_\Gamma\left(-\frac{1}{z}\right) = -z^{n/2}\theta_\Gamma(z),$$

so the differential form $\omega = (dz)^{n/4}\theta_\Gamma(z)$ is sent to

$$\left(d(-\tfrac{1}{z})\right)^{n/4}\theta_\Gamma\left(-\frac{1}{z}\right) = \left(\frac{dz}{z^2}\right)^{n/4} \cdot -z^{n/2}\theta_\Gamma(z) = -\omega$$

under S. But ω is invariant under T, so ST transforms ω into $-\omega$. This is a contradiction with the fact that $(ST)^3$ is the identity, and thus we must have had n be a multiple of 8 from the start. □

Corollary 109

Given any lattice satisfying our assumptions, there exists a cusp form f_Γ of weight $2k = \frac{n}{2}$ such that $\theta_\Gamma = E_k + f_\Gamma$.

Proof. Both θ_Γ and E_k have constant term 1 in their q-expansions, so their difference is a cusp form. □

Taking this result and directly reading off coefficients yields the following:

Corollary 110

For any lattice Γ satisfying our assumptions, we have

$$r_\Gamma(m) = \frac{4k}{B_k}\sigma_{2k-1}(m) + O(m^k),$$

again taking $k = \frac{n}{4}$.

(Here, we need to use that the coefficients $|a_n|$ of a cusp form are $O(m^k)$, and also that k must be even.)

Fact 111

In general, the cusp forms f_Γ are not identically zero. However, Siegel has shown that a weighted mean of the f_Γs is zero: if C_n is the set of isomorphism classes of our Γs, and n_Γ is the order of the automorphism group of Γ in C_n, then

$$\sum_{\Gamma \in C_n} \frac{1}{n_\Gamma} f_\Gamma = 0.$$

Another way to say this is that (because $f_\Gamma = \theta_\Gamma - E_k$)

$$\sum_{\Gamma \in C_n} \frac{1}{n_\Gamma} \theta_\Gamma = \left(\sum_{\Gamma \in C_n} \frac{1}{n_\Gamma} \right) E_k.$$

Since E_k is an eigenfunction of our Hecke operators $T(n)$, so is this weighted mean.

With this, we can look at a few concrete examples.

- The smallest n where a lattice with our desired properties exists is $n = 8$. The theta functions then correspond to modular forms of weight 4, but there are no nonzero cusp forms of weight 4. Thus, we must have

  $$\theta_\Gamma(z) = E_2(z) = 1 + \sum_{m \ge 1} 240\sigma_3(m) q^m$$

  for any 8-dimensional lattice Γ. We have a matrix representation for such a lattice:

  $$\Gamma_8 = \begin{bmatrix} 2 & & -1 & & & & & \\ & 2 & & -1 & & & & \\ -1 & & 2 & -1 & & & & \\ & -1 & -1 & 2 & -1 & & & \\ & & & -1 & 2 & -1 & & \\ & & & & -1 & 2 & -1 & \\ & & & & & -1 & 2 & -1 \\ & & & & & & -1 & 2 \end{bmatrix}$$

  (this has determinant 1), and in fact, this is the only isomorphism class of lattices in C_8.

- The next smallest value of n is $n = 16$. There are still no nonzero cusp forms of weight 8 (for $k = 4$), so we know that

  $$\theta_\Gamma(z) = E_4(z) = 1 + \sum_{m \ge 1} 480\sigma_7(m) q^m.$$

  Notably, this is true even though there are two different isomorphism classes of lattices: Γ_{16} (as defined in Serre chapter 5)[2] and $\Gamma_8 \oplus \Gamma_8$. Since the direct sum of lattices gives a squared generating function, this yields the surprising identity

  $$1 + \sum_{m \ge 1} 480\sigma_7(m) q^m = \left(1 + \sum_{m \ge 1} 240\sigma_3(m) q^m \right)^2.$$

  For example, looking at the $m = 2$ coefficient, we have $1 + 480(2^7 + 1) = 1 + 2 * 240(2^3 + 1) + 240^2$.

- Finally, looking at the case $n = 24$ finally gives us a nonzero cusp form: the space of modular forms is spanned

[2] These lattices are defined by taking a subset of the half-integer lattice points which satisfy certain conditions: for example, the sum of the coordinates must be an even integer.

by the two functions

$$E_6 = 1 + \frac{65520}{691}\sum_{m\geq 1}\sigma_{11}(m)q^m, \quad F = q\prod_{m\geq 1}(1-q^m)^{24} = \prod_{m\geq 1}\tau(n)q^n.$$

Thus, we can write our theta function

$$\theta_\Gamma(z) = E_6 + c_\Gamma F$$

for some constant c_Γ which depends on the lattice. Computing this constant requires us to count the number of points $x \in \Gamma$ with $\langle x, x\rangle = 2$, since computing the q-coefficients yields

$$r_\Gamma(1) = \frac{65520}{691} + c_\Gamma.$$

For example, $\Gamma = \Gamma_8 \oplus \Gamma_8 \oplus \Gamma_8$ has $r_\Gamma(1) = 720$, and $\Gamma = \Gamma_{24}$ has $r_\Gamma(1) = 1104$. Of special attention is the **Leech lattice**, which has applications to coding theory and sphere packings: this particular lattice has $r_\Gamma(1) = 0$, so $c_\Gamma = -\frac{65520}{691}$.

As a closing remark, needing to restrict ourselves to Type II lattices severely restricts the kinds of lattices we can deal with. In particular, forcing that the diagonal terms are even means that we cannot analyze quadratic forms such as

$$x_1^2 + x_2^2 + \cdots + x_n^2.$$

to analyze, for instance, the number of ways in which a positive integer can be written as the sum of n squares. In order to deal with such cases, we relax our conditions slightly: now $\langle x, x\rangle$ can be any integer, so defining an analogous theta function can be done with respect to the **subgroup of G generated by S and T^2.** Such functions have "weight $\frac{n}{2}$" (may not be an integer), and the fundamental domain now has two cusps, yielding two different "Eisenstein series." Further discussion will be postponed to later in this class, though.

9 March 5, 2020

The LMFDB

Today, **David Roe** is here to talk to us about the L-functions and Modular Forms Database. The link to the website can be found at https://www.lmfdb.org.

This project is focused on creating a database of objects in computational number theory – a lot of development is happening at MIT and at other parts of the world. It's open-source, and we can use it to learn about some of the objects we've been talking about in this class!

On the left side of the website, there's a tab with links to various parts of the database. Modular forms, for example, are split up into their four types (classical, Maass, Hilbert, and Bianchi). The database also includes elliptic curves over $\mathbb{Q}$ and over number fields, genus 2 and higher genus curves (calculation gets harder as genus increases), and abelian varieties over finite fields.

Remark 112. *The dimension of the abelian variety attached to the elliptic curve (the* ***Jacobian****) is actually the genus of the curve. We might talk more about this later.*

Other parts of the left column of the website include other objects in algebraic number theory: there's a tab for number fields classified by discriminant (for example, $\mathbb{Q}[\sqrt{d}]$ for d squarefree has discriminant d if $d \equiv 1 \bmod 4$ and

$4d$ otherwise). And we define invariants similarly for other number fields. There's also a classification of Galois groups and Sato-Tate groups (which are related to counting points on elliptic curves mod p).

We'll take a quick look at the **classical modular forms** tab. One nice feature is that there are links which expand out and explain terms: for example, we can click on the word **newform**, and we'll get a definition and further explanation of the mathematics there. So the purpose of the LMFDB is to provide both a research tool for experienced mathematicians and an expository one for those of us who are learning!

This means that there are two ways to look for a particular modular form: first of all, we can browse by the links at the top of the page.

Fact 113

We've only been talking about modular forms of **full level** so far, which means that we want to transform under all matrices in $SL_2(\mathbb{Z})$. But changing the level means we only look at specific subgroups, which gives us more freedom. Modular forms that come up in the proof of Fermat's Last Theorem are of weight 2 and high level.

Secondly, we can search for something more specific. For example, if we know a specific **label** for our modular form (this means we have a permanent URL), we can type it in to find its homepage. Alternatively, we can search for a modular form with specific properties.

So once we find a modular form, we get to its homepage! Every page starts with some parameters and invariants – for example, we are told about the level and weight of the form, its **analytic rank** (the error counts in the point counts for elliptic curves give us a modular form $a_1q + a_2q^2 + a_3q^3 + \cdots$, where a_p is the error term for the elliptic curve under $\mathbb{F}_p$. And we can also put those coefficients into an L-function and look at the order of the zero). We're given the q-expansion – for weight 1, they often live in cyclotomic fields, so we see them in terms of roots of unity.

Fact 114

We can search for modular forms by **dimension** as well, which tells us about the size of the "block" when we try to "diagonalize" the space of modular forms over $\mathbb{Q}$ instead of $\mathbb{C}$. (Each block corresponds to a factor of the characteristic polynomial over $\mathbb{Q}$, and newforms correspond to these blocks.) So a lot of calculations end up being linear algebra here!

It's also good to look at the elliptic curves part of the database, which has also classified in a few nice ways. The discriminant is something we define directly in terms of the equation of the curve, and we also define the **conductor** (which has a more complicated definition, but it has the same prime divisors as the discriminant). We can think of the conductor as the level of the associated modular form!

The individual elliptic curve homepages are interesting too: they tell us about integral points, group structure, certain invariants, and so on. Many pages on this database also have **related pages** linked on the right, which are helpful if we want to learn more.

So how is this database generated? Each section has pages talking about **completeness**, **reliability**, and **source** of the data (for example, there might be assumptions like the generalized Riemann hypothesis).

Writing topics

Professor Kim has updated the Stellar page with sample papers and sample abstracts (from last year's Seminar in Number Theory).

The most important (and maybe most difficult) part of writing a survey paper is to choose a topic, so that's what we'll brainstorm now! Let's (for example) think about **modular forms**. We need to write an **abstract**, which is supposed to (briefly) summarize the main content and results of our paper. For example, the abstract might say that "we will define modular forms and their properties, and we'll prove the dimension formula." It's good to say that "this is a survey/expository paper based on (books/papers)," since we won't really be doing original research.

Most papers start with an **introduction**, which is basically an expanded summary of the content of our paper. But if there is some history or background for the subject, we can also talk about that to give some motivation for the mathematics, and we can also state a main theorem and the ingredients for its proof (if they exist). **One main purpose for the introduction is to state the structure of the paper**! For example, we might say that "in section 2, we do (something), and in section 3, we cover (other things)," summarizing each section.

Remark 115. *When we write an introduction, we might need to introduce new terminology which is not defined yet. In principle, we want to define everything before we use it, but an introduction should not have too many detours. So it's often good to give label numbers to where definitions are actually made!*

Let's think about how we might structure a paper about modular forms, for example. We want to start with a definition – a modular form is a function satisfying a holomorphic condition, as well as invariance with respect to the action of $SL_2(\mathbb{Z})$ – but to have this definition, we need to define things like $\mathbb{H}$, the relations of $SL_2(\mathbb{Z})$, how $SL_2(\mathbb{Z})$ acts on the $\mathbb{H}$, and so on.

Once the definition is established, a good next step is to give some examples (Eisenstein function, discriminant, proving that both of these are indeed modular forms), and then start proving some properties about them. The specific sample paper we can see on Stellar finished by proving the dimension formula, and it also had an **appendix** with results from complex analysis (so that they aren't too distracting).

Fact 116

We should be careful about plagiarism in our papers – we should make sure to quote and cite our sources! We can label sources by numbers or by initials: for example, a book by Diamond and Shurman written in 2005 can be labeled as [DS05] or as [1]. But if we're using a source, we should cite the specific theorem or proof from the book so that it's easy to look up.

Remark 117. *We're writing this paper for an audience like our classmates – basically, we have some background in modular forms.*

10 March 10, 2020

Diamond and Shurman 1.2 – David Wu

We'll start off the new book by talking about **congruence subgroups**. First, a quick review of notation and generalization: let $\gamma \in \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be an element of $SL_2(\mathbb{Z})$. We previously defined what it means for a function to be weakly modular with respect to the full modular group G: basically, f is weakly modular of weight k with respect to $SL_2(\mathbb{Z})$ if

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

for all $\gamma \in SL_2(\mathbb{Z}), \tau \in \mathbb{H}$.

But now we'll make a generalization: we'll replace $SL_2(\mathbb{Z})$ with a **congruence subgroup** Γ, which will allow weights to be any nonnegative integer k.

What is causing this generalization? We know that $-I$ is an element of $SL_2(\mathbb{Z})$, and the group action $-I(\tau) = \tau$ fixes complex numbers $\tau \in \mathbb{H}$. So previously, when k was odd, the definition of a modular function would force

$$f(\tau) = (-1)^k f(\tau),$$

which means that f is trivial. The main idea here is that we won't need to include $-I$ in our congruence subgroup! At the end of the last lecture, we were talking about theta functions, and we introduced (for a lattice Γ)

$$r_\Gamma(m) = \#\{x \in \Gamma : \langle x, x\rangle = 2m\}.$$

We required gnarly conditions on Γ last time to make sure we had a modular form when we assembled the generating function θ_Γ. In particular, we (sort of) showed last time that

$$\theta_\Gamma\left(-\frac{1}{z}\right) = (iz)^{n/2}\theta_\Gamma(z).$$

Let's define this more rigorously now:

Example 118

Let $r(n, k)$ be the number of lattice points $v \in \mathbb{Z}^k$ such that

$$n = \sum_{i=1}^k v_i^2.$$

This defines a **theta function** $\theta(\tau, k) = \sum_{n=0}^\infty r(n, k)q^n$ (where $q = e^{2\pi i\tau}$).

This is mostly as a sketch – we won't show that this actually converges as a power series or anything like that – but the reason this function is important is that we have **Legendre's four squares theorem**, which tells us that any nonnegative integer can be written as the sum of four squares. This can be rephrased as saying that $\theta(\tau, 4)$ has all nonzero coefficients in its q-expansion: once we develop a bit more theory, we'll be able to make progress on this question!

Claim 119. *$\theta(\tau, 4)$ is a modular form of weight 2 with respect to a specific congruence subgroup Γ.*

Proof sketch. First, we can notice that if $i + j = k$, then we have the convolution

$$r(n, k) = \sum_{\ell+m=n} r(\ell, i)r(m, j).$$

Basically, if we want to write n as a sum of k squares, we go through all possible cases of what the first i squares add up to. And this also relates nicely to our generating function: we have

$$\theta(\tau, k_1)\theta(\tau, k_2) = \theta(\tau, k_1 + k_2).$$

Since $q = e^{2\pi i\tau}$, our function θ is $\mathbb{Z}$-periodic. This means that we already have (assuming absolute convergence for now)

$$\theta\left(\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\tau, 4\right) = \theta(\tau + 1) + \theta(\tau, 4)$$

for all θ, so the action under T is "correct." For the action under S, our goal is to show that $\theta\left(-\frac{1}{\tau}, 4\right)$ is related to $\theta(\tau, 4)$. Define the function $\theta = \theta(\tau, 1)$; then we know that $\theta^4 = \theta(\tau, 4)$, so it's good enough to figure out the behavior of θ. θ is the number of ways to write a number as a single perfect square, so $\theta = \sum_{d \in \mathbb{Z}} e^{2\pi i d^2 \tau}$ (possibly with an extra constant factor), and then Poisson summation gives us

$$\theta\left(-\frac{1}{4\tau}\right) = \sqrt{-2i\tau}\theta(\tau).$$

Notice, though, that this transformation gives us the matrix $\begin{bmatrix} 0 & -1 \\ 4 & 0 \end{bmatrix}$, which does not have determinant 1. Instead, we'll have to use

$$\begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1/4 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 4 & 0 \end{bmatrix}$$

which maps τ to $\frac{\tau}{4\tau+1}$, and this gives us (with some algebra)

$$\theta\left(\frac{\tau}{4\tau+1}\right) = \sqrt{4\tau+1}\theta(\tau).$$

This implies (raising everything to the fourth power) that

$$\theta\left(\frac{\tau}{4\tau+1}, 4\right) = (4\tau+1)^2\theta(\tau, 4).$$

□

So now we've verified the modular form equation

$$\theta(\gamma(\tau), 4) = (c\tau + d)^2\theta(\tau, 4)$$

under action of $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix}$. Before, we cared about being weakly modular with respect to S and T, and now we'll look at being weakly modular with respect to these new matrices. But here we're describing things with respect to the actual generators, and we want something more general:

Definition 120

Let N be a positive integer. The **principal congruence subgroup of level N** is

$$\Gamma(N) = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SL_2(\mathbb{Z}) : \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \bmod N \right\},$$

where congruence is taken entry-wise.

As a check, we can make sure $\Gamma(N)$ is actually a subgroup of $SL_2(\mathbb{Z})$ – this can be done just by running through the axioms. Notice that $\Gamma(1) = SL_2(\mathbb{Z})$, and also that $\Gamma(N)$ has **finite index** in $SL_2(\mathbb{Z})$. Later, we'll state a result that helps us explicitly calculate this index, but first we'll finally make our important definitions:

Definition 121

A subgroup $\Gamma \subset SL_2(\mathbb{Z})$ is a **congruence subgroup** if $\Gamma(N) \subset \Gamma$ for some N. Then Γ is called a congruence subgroup of **level N**.

Definition 122

Some special subgroups:

$$\Gamma_0(N) = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SL_2(\mathbb{Z}) : \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} * & * \\ 0 & * \end{bmatrix} \bmod N \right\}$$

and

$$\Gamma_1(N) = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SL_2(\mathbb{Z}) : \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & * \\ 0 & 1 \end{bmatrix} \bmod N \right\},$$

where $*$ can be any element.

This may seem a bit unmotivated, but we'll see why they're useful later. It's possible to find generators explicitly for $\Gamma(N), \Gamma_0(N)$, and $\Gamma_1(N)$, but it's a bit annoying.

Proposition 123

Let N be a positive integer. Then the following results hold:

1. $\Gamma(N) \subset \Gamma_1(N) \subset \Gamma_0(N) \subset SL_2(\mathbb{Z})$.
2. The map $SL_2(\mathbb{Z}) \to SL_2(\mathbb{Z}/n\mathbb{Z})$ (the natural map) is a surjection with kernel $\Gamma(N)$, so $SL_2(\mathbb{Z})/\Gamma(N)$ is isomorphic to $SL_2(\mathbb{Z}/n\mathbb{Z})$.
3. Consider the map $\Gamma_1(\mathbb{Z}) \to \mathbb{Z}/n\mathbb{Z}$, sending $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ to b mod N. Then this map is a surjection with kernel $\Gamma(N)$, so $\Gamma_1(N)/\Gamma(N)$ is isomorphic to $\mathbb{Z}/n\mathbb{Z}$.
4. Consider the map $\Gamma_1(\mathbb{Z}) \to (\mathbb{Z}/n\mathbb{Z})^*$, sending $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ to d mod N. Then this map is a surjection with kernel $\Gamma_1(N)$, so $\Gamma_0(N)/\Gamma_1(N)$ is isomorphic to $(\mathbb{Z}/n\mathbb{Z})^*$.

The main point is that $\Gamma(N), \Gamma_0(N), \Gamma_1(N)$ are not super mysterious: we can find explicit correspondences between them. We'll just do a quick sketch of the second point:

Proof sketch of (2). The identity element of $SL_2(\mathbb{Z})$ is $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, so $\Gamma(N)$ is indeed the kernel here by definition. Showing surjectivity just requires us to lift $SL_2(\mathbb{Z}/n\mathbb{Z})$ to $SL_2(\mathbb{Z})$, and the Euclidean algorithm tells us that this can be done. □

This isomorphism allows us to find explicit formulas for the index of these subgroups, and this tells us that Γ_0 and Γ_1 have finite index (which can be explicitly calculated as well).

Definition 124

Let the **factor of automorphy** $j(\gamma, \tau) = c\tau + d$ for a matrix $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Also, let the **weight-k** operator $[\gamma]_k$ be defined as

$$(f[\gamma]_k)(\tau) = j(\gamma, \tau)^{-k} f(\gamma(\tau)).$$

We find after unpacking definitions that this just gives us a nice way to say that f is weakly modular of weight k if $f[\gamma]_k = f$. We can also translate properties (like those about group actions) that we previously proved into operator notation:

Lemma 125

We have, for all $\tau \in \mathbb{H}, \gamma, \gamma' \in SL_2(\mathbb{Z})$,

- $j(\gamma\gamma', \tau) = j(\gamma, \gamma'(\tau))j(\gamma', \tau)$.
- $(\gamma\gamma')(\tau) = \gamma(\gamma'(\tau))$.
- $[\gamma\gamma']_k = [\gamma]_k[\gamma']_k$ (as operators).
- $\text{Im}(\gamma, \tau) = \frac{\text{Im}\,\tau}{|j(\gamma,\tau)|^2}$.
- $\frac{d\gamma(\tau)}{d\tau} = \frac{1}{j(\gamma,\tau)^2}$.

This is making sure that the action under $\gamma\gamma'$ is the same as the action under γ' followed by the action under γ. The same proofs work – we're just using operator notation, and the idea is that there's a more intrisic view of these operators independent of the modular form f. One nicer way to show these that we haven't discussed yet, though, is to define a group action of $SL_2(\mathbb{Z})$ on column vectors by matrix multiplication. Then we find that

$$\gamma \begin{bmatrix} \tau \\ 1 \end{bmatrix} = \begin{bmatrix} \gamma(\tau) \\ 1 \end{bmatrix} j(\gamma, \tau),$$

and this allows us to show the identities easily.

The point is that if f is weakly modular with respect to a set of matrices, it's also weakly modular with respect to the group that is generated by those matrices. It turns out that the theta function $\theta(\tau, 4)$ is weakly modular of weight 2 with respect to the subgroup $\Gamma_0(4)$.

Diamond and Shurman 1.2 (continued) – Dhruv Rohatgi

We'll now formally define modular forms with respect to the congruence subgroup and apply this to the four square theorem.

Definition 126

Let $\mathcal{M}_k(\Gamma)$ be the vector space of modular forms of weight k over a congruence subgroup Γ. A function $f : \mathbb{H} \to \mathbb{C}$ is **weakly modular of weight *k*** with respect to Γ if

1. f is weakly modular of weight k with respect to Γ,
2. f is holomorphic on $\mathbb{H}$,
3. f is **holomorphic at the cusps**.

Let's define what this last point actually means by thinking about what happened when $\Gamma = SL_2(\mathbb{Z})$. In that case, we knew that $f(z) = f(z+1)$, so f was $\mathbb{Z}$-periodic, and there was a transform $\tilde{f} : D \setminus \{0\} \to \mathbb{C}$ defined by

$$f(z) = \tilde{f}(e^{2\pi i z}) = \tilde{f}(q).$$

Because f was holomorphic on the upper half-plane, $\tilde{f}$ was holomorphic on the punctured unit disk, so there was a Laurent expansion in q: our condition forced us to say that there were no negative coefficients there.

Now, in our general case, we know that $\Gamma \supset \Gamma(N) \ni \begin{bmatrix} 1 & N \\ 0 & 1 \end{bmatrix}$, where this last matrix represents translation by N.

So a function that is weakly modular with respect to Γ satisfies $f(z) = f(z+N)$ for some N, and now we can define

$$f(z) = \tilde{f}(e^{2\pi iz/N}),$$

where we'll define $q = q_N = e^{2\pi iz/N}$. As before, we have a Laurent expansion at q, and we'll say that a function f which just satisfies (1) and (2) (that is, weakly modular of weight k with respect to Γ, and holomorphic at $\mathbb{H}$) is also **holomorphic at ∞** if $\tilde{f}$ has a holomorphic extension at $q = 0$.

But this isn't strong enough for a general subgroup Γ: we need that the vector spaces of modular forms of a fixed weight is finite dimensional. In the $SL_2(\mathbb{Z})$ case, we can map any rational point on the real line to ∞ with a transformation in $SL_2(\mathbb{Z})$, but this isn't quite so nice in the general case:

Definition 127

Let Γ be a congruence subgroup. Then a **cusp** of Γ is an equivalence class of $\mathbb{Q} \cup \{\infty\}$ under action by Γ. (Because Γ has finite index, there are only a finite number of cusps.)

Definition 128

A function $\mathbb{H} \to \mathbb{C}$ satisfying (1) and (2) above is **holomorphic at the cusps** if $f(\gamma z)$ is holomorphic at ∞ for all $\gamma \in SL_2(\mathbb{Z})$.

(Notably, this maps any cusp to infinity.) At first glance, this seems a bit hard to work with: if we're given a q-expansion, it's easy to say that it's holomorphic at ∞, but not so much with other cusps. Fortunately, we have an easier condition to check:

Proposition 129

Let $f : \mathbb{H} \to \mathbb{C}$ be holomorphic (on $\mathbb{H}$) and weakly modular with respect to a congruence subgroup Γ of weight N, and suppose f is holomorphic at ∞ with coefficients of q-expansion

$$f(z) = \sum_{n=0}^{\infty} a_n q^n$$

satisfying $a_n = O(n^r)$ for some constant r (polynomially bounded), then f is modular form with respect to Γ (it is holomorphic at the cusps).

Proof sketch. Doing some calculations shows that if $z = x + iy$, we have

$$|f(z)| \le c + \frac{c_1}{y^{r+1}}$$

by doing some integration. Now if we let $\alpha \in SL_2(\mathbb{Z})$, it suffices to show that $q \cdot f(\alpha z)$ converges to 0 as $q \to 0$. Then $q \cdot f(\alpha z)$ extends to a holomorphic function with a zero, so we can divide out the q again. To do this, we use the above bound: fixing q and taking $q = e^{2\pi iz/N}$ for $x \in [0, N]$ (without loss of generality),

$$\text{Im}(\alpha z) = \frac{\text{Im}\, z}{|cz+d|^2},$$

and using the bound on x and $\text{Im}\, z$, this is at least $\frac{y}{N^2+y^2} \ge \frac{C}{y}$. So the magnitude of $|f(\alpha z)|$ is at most $O(y^{r+1})$, but q is exponentially decaying, so we're done. □

We'll return to the motivating example from David's lecture:

Proposition 130

$\theta_4(z) = \theta(z, 4)$ is a modular form in $\mathcal{M}_2(\Gamma_0(4))$.

Proof. We showed in the previous lecture that $\theta_4(z)$ is weakly modular. To show that this is holomorphic, we know that

$$\theta_4(z) = \sum_{n=0}^{\infty} r(n, 4)q^n.$$

On any compact subset of $\mathbb{H}$, q is bounded by 1, and $r(n, 4)$ is bounded polynomially, so this will converge absolutely. And because $|r(n, 4)| = O(n^4)$, we can use the previous proposition to show that θ_4 is holomorphic at the cusps as well. □

We'll now look at some examples that help us characterize our vector space of modular forms. Now that we can have modular forms of any integer weight, we now write the weight 2 Eisenstein series as

$$G_2(z) = \sum_{n\in\mathbb{Z}} \sideset{}{'}\sum_{m\in\mathbb{Z}} \frac{1}{(nz+m)^2} = 2\zeta(2) - 8\pi^2 \sum_{n=1}^{\infty} \sigma_1(n)q^n.$$

This isn't a modular form in any reasonable sense, but here's how we can use it:

Proposition 131

We have the functional equation

$$G_2(\gamma z) = (cz+d)^2 G_2(z) - 2\pi i c(cz+d)$$

for any $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ in $SL_2(\mathbb{Z})$. Thus, for any positive integer n, we can define

$$G_{2,N}(z) = G_2(z) - NG_2(Nz),$$

and this will be a modular form of weight 2 over $\Gamma_0(N)$.

Proof sketch. We saw in a previous lecture (with some difficulty) that

$$G_z\left(-\frac{1}{z}\right) = z^2 G_2(z) - 2\pi i z,$$

and also that

$$G_2(z+1) = G_2(z).$$

Thus, the functional equation is satisfied for $\gamma = S, S^{-1}, T, -I$ (if we just plug in the correct values of c and d for each). Suppose that $G_2(\gamma_1 z) = (c_1 z + d)^2 G_2(z) - 2\pi i c_1(c_1 z + d_1)$, and also that $G_2(\gamma_2 z) = (c_2 z + d)^2 G_2(z) - 2\pi i c_2(c_2 z + d_2)$. Then we can check by direct computation that $G_2(\gamma_1\gamma_2 z)$ also satisfies the desired equation, and since $S, S^{-1}, T, -I$ generate $SL_2(\mathbb{Z})$, we've shown the functional equation for all γ.

Now if we define $G_{2,N}$ as above, we can show that this is weakly modular with respect to $\Gamma_0(N)$ by direct computation. To show holomorphicity, we do something similar as with θ_4: since the coefficients of $G_2(z)$ and $NG_2(Nz)$ are polynomially bounded and we're subtracting two such terms, their difference is also polynomially bounded. Using the proposition above then shows that $G_{2,N}(z)$ is a modular form. □

Example 132

Consider $N = 2$ in the above calculation, giving us the series $G_{2,2}(N)$.

Then

$$G_{2,2}(z) = G_2(z) - 2G_2(2z),$$

and looking at the q-expansion shows that

$$G_{2,2}(z) = -2\zeta(2) - 8\pi^2 \sum_{n=1}^{\infty} \left(\sigma_1(n)q^n - 2\sigma_1(n)q^{2n}\right) = -2\zeta(2) - 8\pi^2 \sum_{n=1}^{\infty} \left(\sigma_1(n) - 2\sigma_1(n/2)\right) q^n,$$

where σ_1 is zero if its argument is not an argument. Notice that $2\sigma_1(n/2)$ is actually the sum of all even divisors of n, so we can write

$$-2\zeta(2) - 8\pi^2 \sum_{n=1}^{\infty} \left(\sum_{d|n, 2\nmid d} \right) q^n.$$

This is a modular form with respect to $\Gamma_0(2)$.

Example 133

Take $N = 4$, giving us the series $G_{2,4}(N)$.

This, similarly, shows that

$$G_{2,4}(z) = -\pi^2 - 8\pi^2 \sum_{n=1}^{\infty} \left(\sum_{d|n, 4\nmid d} d \right) q^n$$

This is a modular form with respect to $\Gamma_0(4)$. Since $G_{2,2}(z)$ is a modular form with respect to $\Gamma_0(2)$, it's also a modular form with respect to $\Gamma_0(4)$. So now we have three modular forms with respect to $\Gamma_0(4)$, but it turns out (we'll learn in May, perhaps) that the dimension of $\mathcal{M}_2(\Gamma_0(4)) = 2$. This means that θ_4 is in the span of $G_{2,2}$ and $G_{2,4}$ (because those two are linearly independent), and we also know the first few coefficients

$$\theta_4(z) = 1 + 8q + \cdots .$$

This is enough to tell us what linear combination to take: we actually just want

$$\theta_4(z) = -\frac{1}{\pi^2} G_{2,4}.$$

So we know what the q-expansion looks like: we have

$$\theta_4(z) = 1 + 8 \sum_{n=1}^{\infty} \left(\sum_{d|n, 4\nmid d} d \right) q^n.$$

And the coefficients here are always nonzero, because 1 always divides d, and we've proven Lagrange's four squares theorem!

11 March 12, 2020

The two speakers (Kaarel and Anton) for today aren't coming due to sickness, so we'll have to change the schedule a bit. (They'll present after spring break.)

The deadline for our paper abstract will stay the same, but we will have an extension. (We should make sure we do submit it by the end of break, though.)

Office hours, as well as class, will be done by Zoom. We'll be livestreaming student lectures, so the time zone for students is important. We can either use a tablet or use slides / Beamer.

12 March 31, 2020

Diamond and Shurman 1.3 – Kaarel Haenni

We'll start by recalling a lemma from a previous lecture: two bases give the same lattice if and only if we have a basechange matrix $\gamma \in SL_2(\mathbb{Z})$ that gets us from one to the other.

Definition 134

A **complex torus** is a set $\mathbb{C}/\Lambda$ of additive cosets $\{z + \Gamma : z \in \mathbb{C}\}$ with algebraic structure as a quotient of $\mathbb{C}$: addition is given by addition in $\mathbb{C}$, and geometric structure is also induced by $\mathbb{C}$.

This turns out to make $\mathbb{C}/\Lambda$ a **Riemann surface** – we won't define this rigorously, but we can just think of this as things looking "locally like $\mathbb{C}$." A good way to think about $\mathbb{C}/\Lambda$ is that we have a fundamental domain of the lattice Λ, except we identify opposite edges. Topologically, there are neighborhoods that look like normal neighborhoods in $\mathbb{C}$, but there are also neighborhoods that "cross over the boundary:" quotienting one pair of edges gives us a tube, and then quotienting the other pair gives us a torus.

We might have seen this next result from regular complex analysis before:

Theorem 135 (Open mapping theorem for Riemann surfaces)

Let X and Y be two Riemann surfaces, and let $f : X \to Y$ be a holomorphic map. Then either f is constant or f is an open mapping (it maps open subsets to open subsets).

(Here, **holomorphic** means that the map looks holomorphic on neighborhoods that "look like" $\mathbb{C}$.)

Proof sketch. Restrict the map to neighborhoods and then apply the usual open mapping theorem; now take a union over the whole surface. □

Corollary 136

Let X and Y be **compact** Riemann surfaces, and say that $f : X \to Y$ is holomorphic. Then f is either compact or a surjection.

Proof. A theorem of topology says that if X is compact, then $f(X)$ is compact, and therefore $f(X)$ is closed. But X is also the whole space, so X is also open, meaning $f(X)$ is open by the open mapping theorem. Thus $f(X)$ is closed and open, and Y is connected – therefore $Y = f(x) \sqcup \overline{f(x)}$ means that $f(X)$ must be the whole space. □

Now we'll look at a specific set of maps that we care about: in particular, everything from before also applies because complex tori are compact Riemann surfaces.

Proposition 137

Suppose $\phi : \mathbb{C}/\Lambda \to \mathbb{C}/\Lambda'$ is holomorphic. Then there exist $m, b \in \mathbb{C}$ such that $m\Lambda \subset \Lambda'$, and the map has the explicit form $\phi(z+\Lambda) = mz + b + \Lambda'$, where m and b are unique mod Λ'. Additionally, this map is bijective if and only if $m\Lambda = \Lambda'$.

Proof sketch. A lifting theorem from topology means that we can lift ϕ to a map between the universal covering spaces. The universal covering space of a complex torus is $\mathbb{C}$, so we have a map $\tilde{\phi}$ from $\mathbb{C} \to \mathbb{C}$ which satisfies the following commutative diagram:

$$\begin{tikzcd}
\mathbb{C} \arrow[r, "\tilde{\phi}"] \arrow[d, "\pi_\Lambda"] & \mathbb{C} \arrow[d, "\pi_{\Lambda'}"] \\
\mathbb{C}/\Lambda \arrow[r, "\phi"] & \mathbb{C}/\Lambda'
\end{tikzcd}$$

Now for any $\lambda \in \Gamma$, consider the function $f_\lambda(z) = \tilde{\phi}(z+\lambda) - \tilde{\phi}(z)$. Since the diagram commutes, f_λ must map to Λ' ($(z+\lambda)$ and z need to end up in the same element of Λ', because they project to the same thing in our commutative diagram). But Λ' is a discrete subset of $\mathbb{C}$, so any continuous map into Λ' has to be constant. Thus its derivative is zero, which means that $\tilde{\phi}'(z+\Lambda) = \tilde{\phi}(z)'$. But now the derivative of $\tilde{\phi}$ is holomorphic and it's also Λ-periodic – this means that it is bounded, and therefore it is constant by Liouville's theorem. This means that $\tilde{\phi}(z) = mz + b$, as desired, and then finding ϕ from $\tilde{\phi}$ is diagram chasing with the above commutative diagram.

And this is an if and only if statement: any map of this form is holomorphic, as long as $m\Lambda \subseteq \Lambda'$. Explicitly finding two elements that map to the same element (showing that it's not injective) gives us the last part of this theorem. □

We're considering holomorphic maps, which preserves the geometric structure, but we also want to preserve the algebraic structure:

Corollary 138

Let $\phi : \mathbb{C}/\Lambda \to \mathbb{C}/\Lambda'$ be holomorphic, or equivalently $\phi(z+\Lambda) = mz + b + \Lambda'$ with $m\Lambda \subseteq \Lambda'$. Then the following are equivalent:

- ϕ is a group homomorphism.
- $b \in \Lambda'$, which means that $\phi(z+\lambda) = mz + \Lambda'$.
- $\phi(0) = 0$.

This is not too hard to show – we just plug in the properties of the group homomorphism. In particular, this tells us that there exist a nonzero holomorphism from $/CC\Lambda$ to $\mathbb{C}/\Lambda'$ if and only if there is a nonzero $m \in \mathbb{C}$ such that $m\Lambda \subseteq \Lambda'$ (and we have an isomorphism if there is equality).

Example 139

Consider a lattice $\Lambda = \omega_1\mathbb{Z} \oplus \omega_2\mathbb{Z}$ with the usual normalization $\tau = \frac{\omega_1}{\omega_2} \in \mathbb{H}$. Let $\Lambda_\tau = \tau\mathbb{Z} \oplus \mathbb{Z}$, and consider the map

$$\phi_\tau : \mathbb{C}/\Lambda \to \mathbb{C}/\Lambda_\tau$$

given by $\phi_\tau(z+\lambda) = z/\omega_2 + \Lambda$.

We can notice that this indeed maps $\mathbb{C}/\Lambda$ to $\mathbb{C}/\Lambda_\tau$ – this is an isomorphism of complex tori, and in fact this sends our complex tori (up to $SL_2(\mathbb{Z})$) to some $\tau \in \mathbb{H}$. To be more precise, tori are isomorphic if they are sent to the same orbit – the isomorphism classes of complex tori are in bijection with orbits of $SL_2(\mathbb{Z})$ of the upper half plane.

Definition 140

An **isogeny** is a nonzero holomorphic homomorphism between complex tori.

We have the following properties:

- Holomorphic isomorphisms are isogenies.
- Isogenies surject.
- Isogenies have finite kernel. (This is because complex analysis tells us that it has a discrete kernel, and then we have a discrete set in a compact space.)

There are two main examples of isogenies that we care about: these might not be all of the isogenies, and this has to do with **complex multiplication**.

Example 141

The multiply-by-N map $[N] : \mathbb{C}/\Lambda \to \mathbb{C}/\Lambda$ sends a point $z + \Lambda$ to $Nz + \Lambda$.

The kernel here is important: it's the set of points $z + \Lambda$ such that $Nz \in \Lambda$, and we call this the set of ***N*-torsion points**. Letting $E = \mathbb{C}/\Lambda$, we notice that we have a group structure:

$$E[N] = \mathbb{Z}/N\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}.$$

If we draw a picture, this basically means we split up our complex torus into an $N \times N$ grid.

Example 142

Let $C \subseteq E[N]$ be a cyclic group of order N in the N-torsion subgroup. Then the preimages of C in the map $\mathbb{C} \to \mathbb{C}/\Lambda$ form a superlattice of Λ, which we also call C – this is because we can pick an explicit basis (ω_1, ω_2) and also of the generators of C.

Then we get a map $\pi : \mathbb{C}/\Lambda \to \mathbb{C}/C$, where $z + \Lambda$ is sent to $z + C$ (this is a projection map). We can convince ourselves that this is a holomorphic homomorphism, and its kernel is C. Pictorally, the easiest example of this is to take the subgroup generated by $\frac{\omega_2}{N}$: then $\mathbb{C}/C$ looks like a slice of the previous fundamental domain, and then we just project down.

Proposition 143

Any isogeny $\phi : \mathbb{C}/\Lambda \to \mathbb{C}/\Lambda'$ can be factored as

$$\phi : \mathbb{C}/\Lambda \xrightarrow{[n]} \mathbb{C}/\Lambda \xrightarrow{\pi} \mathbb{C}/nK \xrightarrow{\sim} \mathbb{C}/\Lambda',$$

where K is the kernel of ϕ.

Basically we multiply by n, then do a cyclic quotient, and then we do an isomorphism. (Proof omitted for now.)

Proposition 144

$\mathbb{C}/\Lambda$ is **isogenous** to $\mathbb{C}/\Lambda'$ if there exists an isogeny between the two. This is an equivalence relation.

Proof. Being reflexive and transitive are both easy – symmetry is a bit trickier, but the idea is that there is a **dual isogeny** that we can explicitly construct. □

This dual isogeny has a few properties:

- The multiply-by-N map's dual is itself.
- The dual of a cyclic quotient is the dual of a different cyclic quotient of the same order (the "orthogonal one") and then scaling up by N.
- The dual of an isomorphism is its inverse.
- The dual of the isomorphisms is the isomorphisms of the dual, but in the opposite oder (just like functions).
- The dual of the dual is itself.
- The sum of duals is the dual of the sum.

For the purpose that we're using these for, which is studying modular forms, it turns out that isogeny is a better equivalence relation than isomorphism.

The last topic of this section is the Weil pairing, but we don't have too much time to talk about it: recall that

$$E[N] \cong \langle \omega_1/N + \Lambda \rangle \times \langle \omega_2/N + \Lambda \rangle,$$

and let μ_N be the group of Nth roots of unity. We can define an inner product $e_N : E[N] \times E[N] \to \mu_N$ as follows: given $P, Q \in E[N]$, we can assemble a coefficient matrix γ such that

$$\begin{bmatrix} P \\ Q \end{bmatrix} = \gamma \begin{bmatrix} \omega_1/N + \Lambda \\ \omega_2/N + \Lambda \end{bmatrix}.$$

We then define

$$e_N(P, Q) = e^{2\pi i \det \gamma/N},$$

and we can check that this is actually independent of our choice of basis for Λ – this is because determinants are basically the ratio of areas for fundamental cells.

Diamond and Shurman 1.4 – Anton Trygub

Today, we're going to talk about the connection between elliptic curves and complex tori. We'll start with any lattice Λ and consider the **Weierstrass *p* function**

$$\wp(z) = \frac{1}{z^2} + \sum_{\omega \in \Lambda - \{0\}} \left(\frac{1}{(z-w)^2} - \frac{1}{w^2} \right)$$

for all $z \in \mathbb{C}$ not in the lattice Λ. (We'll show later on that this is an absolutely convergent sum.) Notice that that the derivative of this function is

$$\wp'(z) = -2 \sum_{\omega \in \Lambda} \frac{1}{(z-w)^3},$$

and now if we consider any $\omega \in \Lambda$, we can consider the function

$$f(z) = \wp(z+\omega) - \wp(z).$$

Then the derivative is

$$f'(z) = \wp'(z+\omega) - \wp'(z) = 0,$$

because $\wp'$ is periodic, so $f(z)$ is constant. And then we can calculate the exact constant by substituting in $-\frac{\omega}{2}$: then

$$f\left(-\frac{\omega}{2}\right) = \wp\left(\frac{\omega}{2}\right) - \wp\left(-\frac{\omega}{2}\right) = 0,$$

because $\wp$ is an even function. Therefore $\wp(z+\omega) = \wp(z)$ for all $\omega \in \Lambda$, and $\wp$ is periodic.

Definition 145

The Eisenstein series for a lattice Λ are defined as

$$G_k(\Lambda) = \sum_{\omega \in \Lambda - \{0\}} \frac{1}{\omega^k}.$$

We're going to make a connection between elliptic curves with complex tori now through the next few results:

Proposition 146

The Laurent expansion of $\wp(z)$ is

$$\wp(z) = \frac{1}{z^2} + \sum_{n \geq 2,\text{ even}} (n+1) G_{n+2}(\Lambda) z^n$$

for all $0 < |z| < \inf\{|\omega| : \omega \in \Lambda - \{0\}\}$.

Proof. We can rewrite

$$\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2} = \frac{1}{\omega^2}\left(\frac{1}{(1-z/\omega)^2} - 1\right) = \frac{1}{\omega^2}\left(1 + \frac{z}{\omega} + \left(\frac{z}{\omega}\right)^2 + \cdots)^2 - 1\right)$$

and then expanding this out yields

$$= \frac{z}{\omega^3}\left(2 + 3\frac{z}{\omega} + 4\left(\frac{z}{\omega}\right)^2 + \cdots\right).$$

Taking $p = \sup_{\omega \in \Lambda - \{0\}} \frac{|z|}{|\omega|}$ (which is less than 1 by assumption), this sum is bounded absolutely by

$$\left|\frac{z}{\omega^3}\right| (2 + 3p + 4p^2 + \cdots) < \infty.$$

Now summing over ω, since $\sum \frac{1}{|\omega^3|}$ was shown to be absolutely convergent, we do indeed have absolute convergence of this series.

So we can now rearrange terms to find that

$$\wp(z) = \frac{1}{z^2} + \sum_{\omega \in \Lambda - \{0\}} \sum_{n=1}^{\infty} (n+1) \frac{z^n}{\omega^{n+2}},$$

and the terms with odd n cancel because the ωs in the denominator have opposite signs. This yields the desired result by evaluating the inner sum to be $G_{n+2}(\Lambda)$, and now we've shown that $\wp(z)$ is well-defined as well. □

Proposition 147

We have the relation

$$(\wp'(z))^2 = 4\wp(z)^3 - g_2(\Lambda)\wp(z) - g_3(\Lambda),$$

where $g_2(\Lambda) = 60G_4(\Lambda)$ and $g_3(\Lambda) = 140G_6(\Lambda)$.

Proof. From the above proposition, we know that

$$\wp(z) = \frac{1}{z^2} + 3G_4(\Lambda)z^2 + 5G_6(\Lambda)z^4 + O(z^6),$$

and also that

$$\wp'(z) = -\frac{2}{z^3} + 6G_4(\Lambda)z + 20G_6(\Lambda)z^3 + O(z^5).$$

By expansion, we see that $\wp'(z)^2$ and $4\wp(z)^3 - g_2(\Lambda)\wp(z) - g_3(\Lambda)$ are both $\frac{4}{z^6} - \frac{24G_4(\Lambda)}{z^2} - 80G_6(\Lambda) + O(z^2)$. Thus their difference is holomorphic, but it is also Λ-periodic. Thus it is bounded and constant, and taking $z \to 0$ makes this difference go to 0, and thus the difference is zero. □

Proposition 148

Let Λ be a lattice generated by (ω_1, ω_2), and say that $\omega_3 = \omega_1 + \omega_2$. Then the cubic equation satisfied by $\wp(z)$ and $\wp'(z)$ above, $y^2 = 4x^3 - g_2(\Lambda)x - g_3(\lambda)$, can be rewritten as

$$y^2 = 4(x - e_1)(x - e_2)(x - e_3),$$

where $e_i = \wp\left(\frac{w_i}{2}\right)$ are distinct.

Proof. The function $f(z) = \wp(z) - t$ is meromorphic for any t, so there is some translation P of the basis parallelogram such that there are no poles or zeros on that boundary. This boundary also has opposite sides being equal, so that tells us that the contour integral

$$\frac{1}{2\pi i}\int_{\partial P} \frac{f(z)'}{f(z)} = 0.$$

Therefore, the number of poles and zeros inside of ∂P is equal. The only pole is the point inside P in the lattice Λ, and that has order 2, so there are two zeros of f inside P.

But $\wp'$ is an odd function, so $\wp'\left(\frac{\omega}{2}\right) = \wp'\left(-\frac{\omega}{2}\right)$ for any $\omega \in \Lambda$. So then we can take the $e_i = \wp\left(\frac{\omega_i}{2}\right)$ that we defined above, and they will indeed be roots of $4x^3 - g_2(\Lambda) - g_3(\Lambda)$ (because those are exactly the points where $\wp'(z)^2 = 0$). And these roots are pairwise different because $\frac{\omega_i}{2}$ is a double root of $\wp(z) - e_i$ (its derivative is zero at that point too). Since we're only allowed two zeros in P, there is no other value of z in P such that $\wp(z) = e_i$. And this logic works for every point in $\wp(z)$. □

Corollary 149

The function Δ is nonvanishing on $\mathbb{H}$.

Proof. For any $\tau \in \mathbb{H}$, we can consider the corresponding lattice Λ_τ and the cubic $4x^3 - g_2(\Lambda_\tau)x - g_3(\Lambda_\tau)$. This discriminant is $\frac{g_2(\tau)^3}{16} - \frac{27g_3(\tau)^2}{16} = \frac{\Delta(\tau)}{16}$, and this is nonzero because the roots are distinct by the previous proposition. Thus Δ is nonzero. □

So now we know that the map $z \to (\wp_\Lambda(z), \wp_\Lambda(z)')$ takes nonlattice points of $\mathbb{C}$ to points on the elliptic curve $y^2 = 4x^3 - g_2(\Lambda)x - g_3(\Lambda)$. Indeed, for any value of $x \in \mathbb{C}$, we have two values of y that satisfy the equation corresponding to the points (x, y) and $(x, -y)$ respectively. We can also extend this map by defining a point at infinity for the elliptic curve, and this means that we have a **bijection** between each complex torus $\mathbb{C}/\Lambda$ and each elliptic curve of the form $y^2 = 4x^3 - g_2(\Lambda)x - g_3(\Lambda)$.

Call this map $(\wp, \wp')$. $(\wp, \wp')$ actually **transforms the group law** from the complex torus onto the elliptic curve. If we have two points $z_1 + \Lambda, z_2 + \Lambda$ on the torus, then $(\wp(z_1), \wp'(z_1))$ and $(\wp(z_2), \wp'(z_2))$ form a secant (or tangent) line in $\mathbb{C}^2$ of the form $ax + by + c = 0$. Consider the meromorphic function

$$f(z) = a\wp(z) + b\wp'(z) + c.$$

We'll use the argument principle again. When $b \neq 0$, this has a triple pole at 0 and zeros at $z_1 + \Lambda, z_2 + \Lambda$. We must have three zeros in this case, and the third zero must be at the point where $z_1 + z_2 + z_3 = 0$. Meanwhile, when $b = 0$, f has a double pole at 0 but still zeros at $z_1 + \lambda$ and $z_2 + \Lambda$, so we can say that $z_1 + z_2 + z_3 = 0$ in $\mathbb{C}/\Lambda$ as well – this corresponds to the point at infinity $(\wp(0), \wp'(0)$. The whole point is that the points on the elliptic curve that also satisfy $ax + by + c = 0$ are exactly $(\wp(z_i), \wp'(z_i))$, where we've just defined z_3, and thus **collinear triples on the elliptic curve sum to zero in the complex torus**.

We've described a way to go from a complex torus to an elliptic curve, and it turns out we can go in reverse as well:

Theorem 150

For any elliptic curve $y^2 = 4x^3 - a_2x - a_3$ where $a_2^3 - 27a_3^2 \neq 0$, there exists a lattice Λ such that $a_2 = g_2(\Lambda)$ and $a_3 = g_3(\Lambda)$.

Proof. We have the surjective map $j : \mathbb{H} \to \mathbb{C}$, so for any a_2 and a_3, there exists a τ such that

$$j(\tau) = \frac{1728a_2^3}{a_2^3 - 27a_3^2} \implies \frac{g_2(\tau)^3}{g_2(\tau)^3 - 27g_3(\tau)^2} = \frac{a_2^3}{a_2^3 - 27a_3^2}.$$

When a_2 and $a_3 \neq 0$, we can take the reciprocal of both sides and subtract 1 to find that

$$\frac{27g_3(\tau)^2}{g_2(\tau)^3} = \frac{27a_3^2}{a_2^2}.$$

This is an invariant quantity when we scale a lattice by ω: in fact, $g_3(\Lambda) = g_3(\tau)\omega^{-6}$. Pick ω accordingly, and then $g_3(\Lambda) = a_3$ and $g_2(\Lambda) = a_2$ as desired.

In the edge cases, if $a_2 = 0$, then $g_2(\tau) = 0$, and we can just scale ω again to get $g - 3$ to the correct value without worrying about g_2. The same thing works if $a_3 = 0$. □

The important takeaway here is that complex tori and elliptic curves are interchangeable!

13 April 2, 2020

Diamond and Shurman 1.5 – Swapnil Garg

The topic for this section is **moduli spaces and modular curves**, but we'll only be talking about specific examples of the former rather than general theory. Recall that the Weierstrass $\wp$ function gives us a bijection between complex tori (equivalent to fundamental domains of lattices in $\mathbb{C}$) and complex elliptic curves, and the mapping is also a group isomorphism (with addition corresponding to collinearity). We also know that there is a holomorphic group isomorphism (meaning that the two complex tori are isomorphic) if and only if $\Lambda' = m\Lambda$ for some $m \in \mathbb{C}$: this means one is a complex multiple of the other.

The purpose of this lecture is to show that complex elliptic curves are in bijection with orbits of the action of $SL_2(\mathbb{Z})$ on the upper half-plane. We'll define specific kinds of moduli spaces here – these three are the only ones that we're going to be using, so we won't define a moduli space completely. Recall that for any integer N, $E[N]$ is the N-**torsion subgroup** of E, which is the elements in the complex torus with $Nq = 0$. We know that $|E[N]| = N^2$ (we have an N by N grid), and the **Weil pairing** is defined as

$$e_N(P, Q) = e^{2\pi i \det \gamma / N},$$

where $\det \gamma$ is the size of the fundamental domain of the lattice generated by P and Q, relative to the size of Γ/N. A few weeks ago, we defined some congruence subgroups of $SL_2(\mathbb{Z})$:

$$\Gamma_0(N) = \left\{ \gamma \in SL_2(\mathbb{Z}) : \gamma = \begin{bmatrix} * & * \\ 0 & * \end{bmatrix} \mod N \right\},$$

$$\Gamma_1(N) = \left\{ \gamma \in SL_2(\mathbb{Z}) : \gamma = \begin{bmatrix} 1 & * \\ 0 & 1 \end{bmatrix} \mod N \right\},$$

$$\Gamma(N) = \left\{ \gamma \in SL_2(\mathbb{Z}) : \gamma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \mod N \right\}.$$

We'll connect these definitions with our elliptic curves as follows:

Definition 151

An **enhanced elliptic curve for $\Gamma_0(N)$** is an ordered pair (E, C), where E is a complex elliptic curve and C is an order N cyclic subgroup of E. An **enhanced elliptic curve for $\Gamma_1(N)$** is an ordered pair (E, Q), where E is a complex elliptic curve and Q is a point on E of order N. An **enhanced elliptic curve for $\Gamma(N)$** is an ordered pair $(E, (P, Q))$, where E is a complex elliptic curve and (P, Q) generate $E[N]$ with the Weil pairing $e_N(P, Q) = e^{2\pi i/N}$.

Notably, we can generate an enhanced elliptic curve for $\Gamma_0(N)$ from one for $\Gamma_1(N)$ by looking at the cyclic group generated by Q. For all three definitions, we define an equivalence relation such that (E, x) and (E, x') are equivalent if there is an isomorphism that sends $E \to E'$ and $x \to x'$ – we'll denote the equivalence classes as $[E, x]$.

Definition 152

Modding out the sets of enhanced elliptic curves for $\Gamma_0(N), \Gamma_1(N)$, and $\Gamma(N)$ by the equivalence relation yield $S_0(N), S_1(N), S(N)$ respectively, which are examples of **moduli spaces**.

When $N = 1$, all three of these are just the space of complex elliptic curves with equivalence by scaling (that we introduced earlier), since the torsion group is trivial.

Definition 153

The **modular curve** $Y(\Gamma)$ for a congruence subgroup Γ of $SL_2(\mathbb{Z})$ is the quotient space of orbits of $\mathbb{H}$ under the action of Γ, which can also be written as $\Gamma \backslash \mathbb{H}$. Denote $Y_0(N), Y_1(N), Y(N)$ to be the modular curves for $\Gamma_0, \Gamma_1, \Gamma$ respectively.

Recall that Λ_τ denotes the lattice $\tau\mathbb{Z} \oplus \mathbb{Z}$ for any $\tau \in \mathbb{H}$. Such lattices correspond to elliptic curves E_τ, and this is unique up to adding an integer to τ.

Theorem 154

We have the following descriptions for the moduli spaces introduced:

1. The moduli space of $\Gamma_0(N)$ is
$$S_0(N) = \{[E_\tau, \langle 1/N + \Lambda_\tau \rangle] : \tau \in \mathbb{H}\}.$$
2. The moduli space of $\Gamma_1(N)$ is
$$S_0(N) = \{[E_\tau, 1/N + \Lambda_\tau] : \tau \in \mathbb{H}\}.$$
3. The moduli space of $\Gamma(N)$ is
$$S(N) = \{[E_\tau, (\tau/N + \Lambda_\tau, 1/N + \Lambda/\tau)] : \tau \in \mathbb{H}\}.$$

Basically, all of the order-N subgroups are isomorphic to taking the point $1/N$ (which generates a specific subgroup in the lattice). Here, $\frac{1}{N} + \Lambda_\tau$ denotes an additive coset of $\mathbb{C}$, and indeed N times that point is an element of the lattice.

Proof. Note that E is isomorphic to $C/\Lambda_{\tau'}$ for some τ'. The three proofs are similar: for (1), if we have an enhanced elliptic curve (E, P) in $S_0(N)$, we can say that the cyclic subgroup P is generated by $(c\tau' + d)/N + \Lambda_{\tau'}$, and then we can find a, b such that $ad - bc = 1 \bmod N$. In (2), we similarly find a, b for the pount $(c\tau' + d)/N + \Lambda_{\tau'}$, and in (3), we know that $(a\tau' + b)/N + \Lambda_{\tau'}, (c\tau' + d)/N + \Lambda_{\tau'}$ are the two points that generate $E[N]$. In all cases, we get a matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ in $SL_2(\mathbb{Z}/N\mathbb{Z})$, which lifts to an element $\gamma \in SL_2(\mathbb{Z})$. Now the element τ we want is

$$\tau = \gamma\tau' = \frac{a\tau' + b}{c\tau' + d},$$

because we indeed can write this enhanced elliptic curve with $m = c\tau' + d$ and

$$m\Lambda_\tau = m(\tau\mathbb{Z} \oplus \mathbb{Z}) = (a\tau' + b)\mathbb{Z} \oplus (c\tau' + d)\mathbb{Z} = \tau'\mathbb{Z} \oplus \mathbb{Z} = \Lambda_{\tau'},$$

where the third equality comes because we're acting by an element γ^{-1} of $SL_2(\mathbb{Z})$, and then

$$m\left(\frac{1}{N} + \Lambda_\tau\right) = \frac{c\tau' + d}{N} + \Lambda_{\tau'}.$$

In other words, we've found a τ such that our enhanced elliptic curve reduces to the desired form. We now want to show that each equivalence class $[E_\tau, (\text{data})]$ map to equivalent enhanced elliptic curves. This can again be done by writing our elements $\tau = \gamma\tau'$ and verifying that multiplying by $m = c\tau' + d$ sends one to the other. □

Theorem 155

In all three cases of the theorem above, we have equivalences $[E_\tau, (\text{data})] \sim [E_{\tau'}, \text{data}]$ if and only if $\Gamma\tau = \Gamma\tau'$ (so τ, τ' are in the same Γ-orbit and map to the same element in the modular curve). Thus, we have bijections from $S_0(N) \to Y_0(N), S_1(N) \to Y_1(N)$, and $S(N) \to Y(N)$.

Proof. We've shown the backwards direction above already. Now if we have equivalent enhanced elliptic curves from E_τ and $E_{\tau'}$, then $m\Lambda_\tau = \Lambda_{\tau'}$. Since $(m\tau, m)$ now forms a basis for $\Lambda_{\tau'}$, we can write $m\tau = a\tau' + b$ and $m = c\tau' + d$, which gives us $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. We wish to show that γ is indeed in the corresponding congruence subgroup Γ, which we can verify by noting that $m(1/N + \Lambda_\tau)$ goes to $1/N + \Lambda_{\tau'}$. □

We can call these specific representatives **enhanced elliptic curves of the special type**. For example, $(E_\tau, 1/N + \Lambda_\tau)$ is an enhanced elliptic curve of the special type, but not $(E_\tau, -1/N + \Lambda_\tau)$. This theorem indeed tells us that $Y(1) = SL_2(\mathbb{Z})\backslash\mathbb{H}$, which is what we wanted to show initially (because torsion data doesn't tell us anything). We talked about the j-invariant earlier in this class – we can now associate each complex elliptic curve with the orbit $SL_2(\mathbb{Z})\tau$, and can define a value $j(E) = j(\tau)$ for each curve. It turns out that elliptic curves with rational j-values correspond to modular forms – this is connected to Fermat's last theorem!

Since moduli spaces and modular curves are equivalent, we can now take maps of modular curves and make them into maps of moduli spaces.

Example 156

We have a natural map from $Y_1(N) \to Y_0(N)$ taking orbits of τ in $\Gamma_1(N)$ to orbits in $\Gamma_0(N)$. This translates into a map from $S_1(N)$ to $S_0(N)$, which takes $[E, Q]$ to $[E, \langle Q \rangle]$.

Example 157

$\Gamma_1(N)$ is a normal subgroup of $\Gamma_0(N)$, so the quotient acts on $Y_1(N)$. Translating this to our moduli spaces, we get an important map γ, defined as

$$\Gamma_1(N)\gamma : [E, Q] \mapsto [E, dQ]$$

(where d is the bottom-right element of the matrix γ). This is a Hecke operator, and it will come up later.

Recall that modular forms have some weight k by definition: there's a connection here as well.

Definition 158

Let Γ be one of the groups $\Gamma_0(N), \Gamma_1(N), \Gamma(N)$. A function $F : \{$enhanced elliptic curves for $\Gamma \to \mathbb{C}$ is **degree-k homogeneous with respect to Γ** if we have

$$F(\mathbb{C}/m\Lambda, mx) = m^{-k} F(\mathbb{C}/\Lambda, x)$$

for any lattice Λ and complex number m. Here, x is a cyclic group of order N if Γ is $\Gamma_0(N)$, a point of order N if it is $\Gamma_1(N)$, and a pair of points if it is $\Gamma(N)$.

Basically, functions on enhanced elliptic curves give functions on the upper half-plane (as a function of τ).

Definition 159

The **dehomogenized function** $f : \mathbb{H} \to \mathbb{C}$ corresponding to a degree-k homogeneous F is

$$f(\tau) = F(\mathbb{C}/\Lambda_\tau, x),$$

where x is $\langle 1/N + \Lambda_\tau \rangle$ if $\Gamma = \Gamma_0(N)$, $1/N + \Lambda_\tau$ if $\Gamma = \Gamma_1(N)$, and $(\tau/N + \Lambda_\tau, 1/N + \Lambda_\tau)$ if it is $\Gamma(N)$.

Proposition 160

A degree-k dehomogenized function f is weight-k invariant with respect to Γ. In other words, $f(\gamma(\tau)) = (c\tau + d)^k f(\tau)$.

We can see this by letting $m = (c\tau + d)^{-1}$ in the definition.

Recall that lattice functions correspond to functions on the upper half-plane: we're saying here that functions on enhanced elliptic curves do the same, and we can get a function F on enhanced elliptic curves **from a function** f that is weight-k invariant with respect to a congruence subgroup. This is well-defined, because two enhanced elliptic curves of the special type will have the same values of f, and then we can define

$$F(\mathbb{C}/\Lambda_{\tau'}, x) = m^{-k} F(\mathbb{C}/\Lambda_\tau, x).$$

Diamond and Shurman 2.1-2 – Natalie Stewart

The point of this lecture is to start allowing us to talk about everything from a differential geometry perspective with Riemann surfaces. Recall that a congruence subgroup $\Gamma \subseteq SL_2(\mathbb{Z})$ is a subgroup containing one of the principal congruence subgroups $\Gamma(N)$ – then the modular curve for Γ is the orbit space $Y(\Gamma) = \Gamma \mathbb{H}$. But we only really know how this looks with respect to the quotient topology, and we're going to upgrade this now.

Definition 161

A **Riemann surface** is a connected complex 1-dimensional manifold – that is, it's a connected topological space with a countable basis which is Hausdorff, and for any two intersecting neighborhoods U_m, U_n, we have **coordinate charts** $\phi_m, \phi_n \to D$ (the unit disk) such that the following composition is holomorphic:

$$\phi_m(U_m \cap U_n) \xrightarrow{\sim} U_m \cap U_n \xrightarrow{\sim} \phi_n(U_m \cap U_n).$$

One note about the last point is that we can map U_m and U_n conformally to the unit disk D, and we want the "transfer map" between the images of the intersection $U_m \cap U_n$ onto the disks to be holomorphic.

We have a surjective open mapping $\mathbb{H} \to Y(\Gamma)$ (which maps τ to $\Gamma\tau$), which tells us immediately that $Y(\Gamma)$ is second-countable and connected. To show that $Y(\Gamma)$ is a Riemann surface, we need to show that it is Hausdorff, providing coordinate charts on $Y(\Gamma)$, and we'll show that those transfer maps defined above are holomorphic.

Proposition 162

Let τ_1, τ_2 be two points in the upper half-plane. Then the action of $SL_2(\mathbb{Z})$ (and of Γ) is **properly discontinuous**: we can find neighborhoods U_1 and U_2 of τ_1, τ_2 such that

$$\gamma(U_1) \cap U_2 \neq \varnothing \implies \gamma(\tau_1) = \tau_2.$$

This takes a long time to prove, so we won't talk too much about this here.

Corollary 163

The modular curve $Y(\Gamma)$ is Hausdorff for any congruence subgroup Γ.

Proof. Suppose we have two different points $\pi(\tau_1), \pi(\tau_2) \in Y(\Gamma)$. By definition, τ_1, τ_2 are in different orbits, so we can choose two neighborhoods as in the above proposition so that $\gamma(U_1) \cap U_2 = \varnothing$, which means that $\pi(U_1) \cap \pi(U_2) = \varnothing$

(all of the points in the neighborhood of τ_1 are in different orbits as those in τ_2). Since π is a quotient mapping, it's an open mapping, and thus $\pi(U_i)$ is a neighborhood of $\pi(\tau_i)$, which gives us our desired neighborhoods. □

Giving coordinate charts is significantly more difficult, so we'll go through some random-seeming constructions which will come together:

Definition 164

Let G be a group acting on a space X, and let $\Gamma \subset G$ be a subgroup. The **isotropy subgroup for τ under Γ** (for any $\tau \in X$) is the set of $\gamma \in \Gamma$ which fix τ.

Definition 165

A point $\tau \in \mathbb{H}$ is an **elliptic point for Γ** if the containment $\{\pm I\}\Gamma_\pi \supset \{\pm I\}$ is a proper containment. ($\pi(\tau)$ is also called elliptic.)

In other words, elliptic points are preserved by some nontrivial transformation of $\mathbb{H}$ by Γ.

Proposition 166

Let Γ be a congruence subgroup of $SL_2(\mathbb{Z})$. Then the isotropy subgroups Γ_τ are finite cyclic groups for all elliptic points τ.

We'll prove this next lecture – this was also proved in Serre. This allows us to make the following definition:

Definition 167

The **period** of $\tau \in \mathbb{H}$, denoted h_τ, is

$$h_\tau = [\{\pm I\}\Gamma_\tau / \{\pm I\}] = \begin{cases} |\Gamma_\tau|/2 & -I \in \Gamma_\tau, \\ |\Gamma_\tau| & -I \notin \Gamma_\tau. \end{cases}$$

Using a bit of general topology, we'll define this for the modular curve as well:

Lemma 168

Let G be a group acting on a space X, and let $\Gamma \subset G$. If $\tau \in X$ is a point, then for any $\gamma \in G$, we have

$$\gamma(\Gamma_\tau)\gamma^{-1} = (\gamma\Gamma\gamma^{-1})_{\gamma(\tau)}.$$

When $\gamma \in \Gamma$, this reduces to the period of τ being constant across orbits – thus, we can associate the period with points of $Y(\Gamma)$ instead of τ.

Proof sketch. If we take any $\alpha \in \Gamma_\tau$, we know that $\gamma\alpha\gamma^{-1}\gamma(\tau) = \gamma\alpha(\tau) = \gamma(\tau)$, which shows that the left side is contained in the right. The other direction is similar – if $\gamma\alpha\gamma$ is an element of the right side, then $\gamma\alpha(\tau) = \gamma(\tau)$, so $\alpha(\tau) = \tau$. □

We'll now move to something that looks unrelated: we can extend an action of $SL_2(\mathbb{R})$ on $\mathbb{H}$ to an action of $GL_2(\mathbb{C})$ on the Riemann sphere:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} z = \frac{az+b}{cz+d}.$$

Definition 169

The **straightening map** for a point $\tau \in \mathbb{H}$ is the action of $\begin{bmatrix} 1 & -\tau \\ 1 & -\overline{\tau} \end{bmatrix} \in GL_2(\mathbb{C})$.

Note that $\delta_\tau(\tau) = 0$ and $\delta_\tau(\overline{\tau}) = \infty$, which is nice because we can map neighborhoods to neighborhoods of 0. It turns out this is also a straightening in a more powerful sense if we connect this with isotropy groups: since congruence subgroups have all real coordinates, $\Gamma_\tau = \Gamma_{\overline{\tau}}$, which helps when we're dealing with the Riemann sphere (and complex numbers). Using the above lemma about conjugation, we find an equality of isotropy groups

$$(\delta_\tau(\{\pm I\}\Gamma)\delta_\tau^{-1})_0/\{\pm I\} = \delta_\tau(\{\pm I\}\Gamma_\tau/\{\pm I\})\delta_\tau^{-1} = (\delta_\tau(\{\pm I\}\Gamma)\delta_\tau^{-1})_\infty/\{\pm I\},$$

and we'll denote this group G. (In words, the transformations induced by the isotropy group of the image of Γ under conjugation of the straightening map at 0 and ∞ are the same as the conjugate of the isotropy groups around τ.) We know that G is given by the properties of the isotropy groups, and for all $g \in G$, we know that 0 and ∞ are preserved, so we must actually have $g(z) = az$. In addition, because G is finite cyclic, each element of G is just a rotation around the origin by some integer multiple of $\frac{2\pi}{h_\tau}$, since $|G| = h_\tau$.

This helps us think about how to define our coordinate charts using our straightening maps: under the map δ_τ, neighborhoods $\pi(U)$ turn into radial sectors of the neighborhood of 0. Then we can apply the h_τ-fold "wrapping map," $\rho_{h_\tau}(z) = z^{h_\tau}$, to get the whole neighborhood.

So we need to decide what our small neighborhood U looks like, and we'll use the following corollary of proper discontinuity of $SL_2(\mathbb{Z})$ action:

Corollary 170

Let Γ be a congruence subgroup of $SL_2(\mathbb{Z})$. Then $\tau \in \mathbb{H}$ has a neighborhood U such that

$$\gamma(U) \cap U \neq \varnothing \implies \gamma \in \Gamma_\tau.$$

This neighborhood has no elliptic points except possibly τ.

Now, we can pick a point $\pi(\tau) \in Y(\Gamma)$, and we'll let U be the neighborhood of τ from the above corollary. We'll introduce the notation $\delta = \delta_\tau, \rho = \rho_{h_\tau}$ (for the straightening map and wrapping map, respectively), $\psi = \rho \circ \delta$, and $V = \psi(U)$.

We claim there is a **bijection** $\phi : \pi(U) \to V$ such that $\phi \circ \pi = \psi$. This is essentially an algebraic manipulation: $\pi(\tau_1) = \pi(\tau_2)$ if and only if the two points are in the same orbit, which is true if and only if τ_1 is in $\Gamma_\tau \tau_2$. This means that

$$\delta(\tau_1) \in \delta(\Gamma_\tau \tau_2) \implies \delta(\tau_1) \in (\delta \Gamma_\tau \delta^{-1})(\delta(\tau_2)),$$

which means that the two elements have the same order – $\delta(\tau_1)^h = \delta(\tau_2)^h$, meaning that $\psi(\tau_1) = \psi(\tau_2)$.

Lemma 171

ϕ is a homeomorphism.

Proof. This is a verification that the map and its inverse are continuous: we know that ψ and π are both continuous open surjections, an open subset W of V will correspond to preimages that are open. □

We now need to show that ϕ is holomorphic, and this is a lot harder. We'll define

$$V_{1,2} = \phi_1(\pi(U_1) \cap \pi(U_2))$$

(the small subset of the disks corresponding to the intersection of the neighborhoods), and define $V_{2,1}$ similarly.

Proposition 172

The transition map $\phi_{2,1} : V_{1,2} \to V_{2,1}$, defined as $\phi_2 \circ \phi_1^{-1}$, is holomorphic.

Proof. Fix a point x in the intersection of the two neighborhoods, and choose some preimages such that $x = \pi(\tau_1) = \pi(\tau_2)$. Fix γ such that $\tau_2 = \gamma(\tau_1)$ (they're in the same orbit under τ, so this exists). We'll check holomorphicity locally on $\phi_1(x)$ – specifically, if $U_{1,2} = U_1 \cap \gamma^{-1}(U_2)$, we'll prove holomorphicity on $\phi_{2,1} = \phi_1 \circ \pi(U_{1,2})$.

Define $\delta_i = \delta_{\tau_i}$ and $h_i = h_{\tau_i}$ for convenience. First assume that $\phi_1(x) = 0$. We know that any point $q = \phi_1(x')$ in this neighborhood in question ($\phi_1 \circ \pi(U_{1,2})$) is of the form

$$q = \psi_1(\tau') = (\delta_1(\tau'))^{h_1}$$

for some $\tau' \in U_{1,2}$, where $\psi_1 = \phi_1 \circ \pi$. Let $\tilde{h_2}$ be the period of the point $\tilde{\tau_2} \in U_2$ – this $\tilde{\tau_2}$ is the point such that $\psi_2(\tilde{\tau_2})$ maps to 0. Then we can chase some equalities to find that

$$\phi_{2,1}(q) = \psi_2(\gamma(\tau')) = \left((\delta_2\gamma\delta_1^{-1})(q^{1/h_1})\right)^{\tilde{h_2}}.$$

The central trick here is that we stick in a $\delta_1^{-1}\delta_1$ term to introduce a q. We know that $\delta_2\gamma\delta_1^{-1}$ is a fractional linear transformation, and we can note that 0 and ∞ are both fixed. Thus, $\delta_2\gamma\delta_1^{-1}$ is just multiplication by some a: we now have a nice form for the transfer map

$$\phi_{2,1}(q) = (aq^{1/h_1})^{\tilde{h_2}}.$$

If $h_1 = 1$, this is holomorphic. Otherwise, τ_1 is elliptic – $\tau_2 = \gamma(\tau_1)$ is in the same orbit, so it is elliptic with the same period. We defined U_2 in such a way that there is only one elliptic point, so $\tau_2 = \tilde{\tau_2}$, so $h_1 = \tilde{h_2}$. But this is exactly what we want: this means that $\phi_{2,1}(q) = a^{\tilde{h_2}q}$, which is a holomorphic map as desired.

We've only proved this for the case where $\phi_1(x) = 0$, but this also proves it for the case where $\phi_2(x) = 0$ (the inverse of a bijective, holomorphic map is holomorphic). And now we can add an intermediate function: write $\phi_{2,1}$ as $\phi_{2,3} \circ \phi_{3,1}$, where $\phi_3 : U_3 \to V_3$ sends our point x to 0, and the composition of two holomorphic maps is holomorphic. □

And this tells us that modular curves $Y(\Gamma)$ are Riemann surfaces, which is exactly what we wanted to show.

14 April 7, 2020

Diamond and Shurman 2.3 – Andrew Gu

Recall that an **elliptic point** for a congruence subgroup Γ is a point in $\tau \in \mathbb{H}$ such that

$$\Gamma_\tau = \{\gamma \in \Gamma : \gamma\tau = \tau\}$$

is larger than $\{\pm I\}$. We'll also call the image of τ in the modular curve $Y(\tau)$ elliptic, and we'll say that the **period** of such an elliptic curve is $|\Gamma_\tau|/2$ when $-I \in \Gamma/\tau$ and $|\Gamma_\tau|$ otherwise.

Last time, we skipped the proof that Γ_τ is a finite cyclic group, and that's going to be the main purpose of this lecture. It turns out that most of the proof is showing that this works for $Y(1)$, so we'll start by looking at the elliptic points there.

Recall that the **fundamental domain** of the upper half-plane is defined as

$$\mathcal{D} = \left\{\tau \in \mathbb{H} : |\operatorname{Re}(\tau)| \le \frac{1}{2}, |\tau| \ge 1\right\},$$

and the generators of $SL_2(\mathbb{Z})$ are

$$S = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

We saw that the map $\pi : \mathbb{D} \to Y(1)$ is surjective in one of the first lectures: basically we translate with T and then use S to get the imaginary part large enough. However, this map is not injective: we know that if $\pi(\tau_1) = \pi(\tau_2)$, then either we're working with the boundary points on the left and right or with the unit circle. The idea is that for $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$,

$$\tau_2 = \gamma\tau_1 \implies |c\tau_1 + d| \le 1.$$

When $c = 0$, we know that $\gamma = \pm \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix}$, which means $\tau_2 = \tau_1 + b$. On the other hand, when $c = \pm 1$,

$$(\operatorname{Re}(\tau_1) + d)^2 \le 1 - \operatorname{Im}(\tau_1)^2 \le \frac{1}{4},$$

which means $|d| \le 1$ and we can again characterize the elliptic points. This gives us the result:

Proposition 173

The elliptic points of $SL_2(\mathbb{Z})$ are $SL_2(\mathbb{Z})i$ and $SL_2(\mathbb{Z})\omega$, where $\omega = e^{2\pi i/3}$. Thus, the isotropy groups for elliptic curves are conjugates of $\langle S \rangle$ and $\langle ST \rangle$, which have order 4 and 6 respectively. Therefore, all elements of finite order are conjugate to one of $-I, S^{\pm 1}, (ST)^{\pm 1}, (STST)^{\pm 1}$, which have order $2, 4, 6, 3$.

Note that $\tau = i$ and $\tau = \omega$ both generate lattices $\mathbb{Z} + \mathbb{Z}\tau$ – one is square and one is hexagonal. Both of these have rotational automorphisms around the origin, which tells us more about the complex elliptic curves associated to the lattices. But we don't have too much to say about this right now.

With this, we can think about the elliptic points of $Y(\Gamma)$ in general:

Proposition 174

Let Γ be a congruence subgruop. Then $Y(\Gamma)$ always has finitely many elliptic points, each with finite cyclic isotropy group and period 2 or 3.

Proof. Pick any $\tau \in \mathbb{H}$. The isotropy group Γ_τ is a subgroup of $SL_2(\mathbb{Z})_\tau$. Since $SL_2(\mathbb{Z})_\tau$ is always cyclic from the above argument, so is Γ_τ. We can think about what possible groups the Γ_τ can be: it can be a subgroup of $\langle S \rangle$, and in this case it can't be a proper subgroup because we're excluding $\{\pm I\}$, so we must be conjugate to $\langle S \rangle$ itself. Thus the group has order 4, meaning that the elliptic point has order $\frac{4}{2} = 2$. Similarly, if $SL_2(\mathbb{Z})_\tau$ has order 6, then the elliptic point has order 3.

To show that there are finitely many elliptic points, note that Γ has finite index in $SL_2(\mathbb{Z})$, so we can write it as a union of cosets $\Gamma\gamma_j$. We know that the elliptic points for Γ are always images of i or ω, so they are a subset of

$SL_2(\mathbb{Z})i \cup SL_2(\mathbb{Z})\omega$. Taking the images in $Y(\gamma)$ means that our elliptic points are a subset of

$$E_\gamma = \{\Gamma\gamma_j i, \Gamma\gamma_j\omega : 1 \le j \le d\},$$

which is at most $2d$ elliptic points, as desired. □

In general, we don't have exactly $2d$ elliptic points: recall that we worked with $\Gamma(N), \Gamma_1(N), \Gamma_0(N)$ in previous lectures.

Proposition 175

$\Gamma(N)$ for $N > 1$, $\Gamma_1(N)$ for $N > 3$, and $\Gamma_0(N)$ for N divisible by a prime $p \equiv -1 \bmod 12$ all have no elliptic points.

We know that a nontrivial isotropy group must be generated by an element of order $3, 4, 6$, so they must have characteristic polynomial $x^2 + 1$ or $x^2 \pm x + 1$. The idea of the proof is to show that there are no generators of this form in the subgroup.

Proof. For $\Gamma(N)$, note that it is a normal subgroup of $SL_2(\mathbb{Z})$, so conjugates of its elements are also in $\Gamma(N)$. And notice that $S, T, ST, (ST)^2$ are all not in $\Gamma(N)$, so there are no elements of order $3, 4, 6$.

For $\Gamma_1(N)$, notice that the trace of the matrix is 2 mod N, and the trace of matrices of order $3, 4, 6$ are all $-1, 0, 1$. So whenever $N > 3$, none of those numbers are 2 mod N.

Finally, for $\Gamma_0(N)$, suppose that there is a prime $p \equiv -1 \bmod 12$ that divides N. We have a matrix $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ with $c \equiv 0 \bmod p$, and we know that the trace is $a + d \in \{-1, 0, 1\} \bmod p$, while $\det\gamma = ad \equiv 1 \bmod p$. Having integer solutions to these equations means that $t^2 - 4ad$ must be a perfect square in $\mathbb{Z}/p\mathbb{Z}$. But whenever p is $-1 \bmod 4$,

$$\left(\frac{-4}{p}\right) = \left(\frac{-1}{p}\right) = -1,$$

and similarly when p is $-1 \bmod 3$, -3 is not a quadratic residue. Thus we get the result that we want. □

For Γ_1, it turns out that $N = 2, 3$ both have elliptic curves: $\Gamma_1(2)$ contains the element $\begin{bmatrix} 1 & -1 \\ 2 & -1 \end{bmatrix}$, and we can solve the equation $\tau = \frac{\tau - 1}{2\tau - 1}$ to get the elliptic point $\frac{1}{2} + \frac{i}{2}$ with period 2. Similarly, $\Gamma_1(3)$ contains $\begin{bmatrix} 1 & -1 \\ 3 & -2 \end{bmatrix}$, and $\frac{1}{2} + \frac{\sqrt{3}}{6}i$ is our elliptic point in this case.

We can use this to classify elliptic points of $\Gamma_0(N)$. Looking ahead, ε_2 and ε_3 count the number of elliptic points of order 2 and 3, and they show up in the dimension formulas for modular functions.

Proposition 176

The period 2 elliptic points of $\Gamma_0(N)$ are in bijective correspondence with ideals J of $\mathbb{Z}[i]$ such that $\mathbb{Z}[i]/J \cong \mathbb{Z}/n\mathbb{Z}$. Similarly, the period 3 elliptic points of $\Gamma_0(N)$ are in bijective correspondence with ideals J of $\mathbb{Z}[e^{2\pi i/6}]$ such that $\mathbb{Z}[e^{2\pi i/6}]/J = \mathbb{Z}/n\mathbb{Z}$.

Basically, we consider elements of order $3, 4, 6$ and look at quadratic residues again. This yields the following formulas:

Proposition 177

The number of elliptic points of order 2 in $\Gamma_0(N)$ is $\prod_{p|N}\left(1+\left(\frac{-1}{p}\right)\right)$ when $N \not\equiv 0 \bmod 4$, and 0 otherwise. Similarly, the number of elliptic points of order 3 in $\Gamma_0(N)$ is $\prod_{p|N}\left(1+\left(\frac{-3}{p}\right)\right)$ when $N \not\equiv 0 \bmod 9$, and 0 otherwise. Here, we're using the Legendre symbol except with the convention that $\left(\frac{-1}{2}\right) = \left(\frac{-3}{3}\right) = 0$ – this is only because $p = 2$ is weird when we're solving the equation $n^2 + 1 \equiv 0 \bmod N$.

Diamond and Shurman 2.4-5 – Nikhil Reddy

We'll talk about **cusps** in this lecture. Recall from previous lectures that a congruence subgroup $\Gamma(N)$ provides a left action on the upper half-plane, and then we define the modular curve $Y(\Gamma)$ to be the Γ-equivalence classes on $\mathbb{H}$. Last time, we showed that $Y(\Gamma)$ is a Riemann surface, and we'll be extending that result here.

If we look at the fundamental domain $\mathcal{D}$ (which is the points with real part between $-\frac{1}{2}$ and $\frac{1}{2}$, outside of the unit circle) on the Riemann sphere, the region looks like a triangle (circles and lines are basically equivalent), except that we're missing the point at infinity. So we'll introduce that now: when we adjoin it, we also need to adjoin its orbit in $SL_2(\mathbb{Z})$. We can show that the orbit is rational, but we need to first show that we can get frmo ∞ to any rational point:

Lemma 178

For any $r = \frac{a}{b} \in \mathbb{Q}$, there exists $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SL_2(\mathbb{Z})$ such that γ is sent to r.

To prove this, we just find $c, d \in \mathbb{Z}$ such that $ad - bc = 1$ by the Chinese Remainder Theorem.

Definition 179

Define $\mathbb{H}^*$ to be $\mathbb{H} \cup \mathbb{Q} \cup \{\infty\}$. The **compactified modular curve** $X(\Gamma)$ is

$$X(\Gamma) = \Gamma\backslash\mathcal{H}^* = Y(\Gamma) \cup \Gamma\backslash(\mathbb{Q} \cup \{\infty\}).$$

We use the notation $X_0(N), X_1(N), X(N)$ for the special congruence subgroups. A **cusp** is a Γ-equivalence class of $\Gamma\backslash(\mathbb{Q} \cup \{\infty\})$.

Proposition 180

$X(\Gamma)$ always has finitely many cusps. In particular, $X(1)$ has only one cusp as we showed above.

Proof. The index of Γ is finite in $SL_2(\mathbb{Z})$, and now take all of the left coset representatives and multiply cusps by them to generate other Γ equivalence classes. This reaches every rational point, so we must have reached all of the equivalence classes, and thus there are finitely many cusps in total. □

We use the Euclidean topology for $\mathbb{H}$, but we need to do a bit more for $\mathbb{H}^*$.

Definition 181

A **neighborhood around** ∞ is of the form

$$\mathcal{N}_M = \{\tau \in \mathbb{H} : \mathrm{Im}(\tau) > M\}.$$

The idea here is that circles centered at ∞ look like lines. Then we can define neighborhoods $\mathcal{N}_M \cup \{\infty\}$, and we can also get neighborhoods around rationals by defining $\alpha(\mathcal{N}_M \cup \{\infty\})$ to be a neighborhood around $\alpha(\infty)$ for any $\alpha \in SL_2(\mathbb{Z})$. Since we've defined our neighborhoods this way, γ is always a homeomorphism, and we get the **quotient topology** from the natural projection map $\pi : \mathcal{H}^* \to X(\Gamma)$.

Basically, the neighborhoods around rational points are open disks tangent at the rational point, plus the rational point itself.

Theorem 182

$X(\Gamma)$ is a compact Riemann surface.

Recall that a Riemann surface needs to be connected, Hausdorff, and it needs a series of charts (sometimes called an atlas), which are pairs (U_m, V_m) such that U_m is a neighborhood around m and V_m is some open set in $\mathbb{C}$ (often the open disk) such that we have a homeomorphism $\phi_m : U_m \to V_m$. In addition, if U_m and U_n intersect, the induced map

$$\phi_m(U_m \cap U_n) \to \phi_n(U_m \cap U_n)$$

should be holomorphic as a function from $\mathbb{C}$ to $\mathbb{C}$. And we just need to add being compact to this list.

Proposition 183

$X(\Gamma)$, the compactified modular group, is connected, compact, and Hausdorff.

Proof. First, we show that $\mathbb{H}^*$ is connected and compact, and then we can project down with the projection map. We showed previously that $\mathbb{H}$ is connected, so if $\mathbb{H}^*$ is a union of two disjoint open sets, then one must contain $\mathbb{H}$ and the other is contained in $\mathbb{Q} \cup \{\infty\}$. But there aren't any nontrivial open sets of $\mathbb{Q} \cup \{\infty\}$.

For compactness, we can first show that the fundamental domain plus the point at infinity $\mathcal{D} \cup \{\infty\}$ is compact in $\mathbb{H}^*$. Then by definition, we know that translates of $\mathcal{D}^*$ cover $\mathbb{H}^*$, so

$$\mathbb{H}^* = SL_2(\mathbb{Z})(\mathcal{D}^*),$$

and then we can break up $SL_2(\mathbb{Z})$ into the right cosets of Γ:

$$= \bigcup_j \Gamma\gamma_j(\mathcal{D}^*).$$

Apply π to each term to find that

$$X(\Gamma) = \bigcup_j \pi(\Gamma\gamma_j(\mathcal{D}^*)).$$

But π and γ_j are both continuous maps, so this means $X(\Gamma)$ is a finite union of compact sets, which means it is compact. (We have a finite union because Γ has finite order in $SL_2(\mathbb{Z})$.)

Being Hausdorff is a bit more complicated: recall that the idea is that two points p_1, p_2 must have neighborhoods U_1, U_2 such that $U_1 \cap U_2 = \varnothing$. We now only need to consider the cases where the two points aren't both in $\mathbb{H}$ (because we did that in the previous class).

In the first case, when $s_1 \in \mathbb{Q} \cup \{\infty\}$ and $\tau_2 \in \mathbb{H}$, take some neighborhood U_2 of τ_2. We need to show that there is a neighborhood of s_1 that does not intersect γU_2 for any $\gamma \in \Gamma$, but note that

$$\text{Im}(\alpha(\tau)) \le \max\{\}.$$

Thus we get an absolute upper bound on the imaginary part of $\gamma(\tau)$, which means that we just need to pick a large enough M such that $\mathcal{N}_M \cup \{\infty\}$ does not intersect $SL_2(\mathbb{Z})U_2$. And now if $\alpha_1(\infty) = s_1$, we have

$$SL_2(\mathbb{Z})U_2 \cap \alpha_1(\mathcal{N}_M \cup \{\infty\}) = \varnothing,$$

and we've found our two neighborhoods.

In the other case where we have two points $s_1, s_2 \in \mathbb{Q} \cup \{\infty\}$, we'll pick two neighborhoods (such that α_1 takes ∞ to s_1 and α_2 takes ∞ to s_2)

$$U_1 = \alpha_1(\mathcal{N}_2 \cup \{\infty\}), \quad U_2 = \alpha_2(\mathcal{N}_2 \cup \{\infty\}).$$

In $\mathcal{N}_2$, we're high enough in the imaginary axis that the only way we can have equivalence is a translation. If the neighborhoods intersect, then we know that $\gamma\alpha_1(\tau_1) = \alpha_2(\tau_2)$ for some $\gamma \in \Gamma$ (the two points are equal up to a γ-action). This means that $\alpha_2^{-1}\gamma\alpha_1$ is a translation, which means that $\gamma(s_1) = s_2$ because ∞ is fixed. But that means $s_1 = s_2$. Thus the neighborhoods don't intersect for distinct points, and now we've indeed showed that $X(\Gamma)$ is Hausdorff. □

From here, we have to deal with our charts: last week, we discussed this for any $\tau \in \mathbb{H}$. Basically, we start with some open set U whose center is an elliptic point – because things look equivalent around such elliptic points, we'll have regions that are Γ-equivalent, which means its map to a unit disk is not so nice. So we had to use a **straightening map** $\delta = \begin{bmatrix} 1 & -\tau \\ 1 & -\overline{\tau} \end{bmatrix}$ last time, which sends our neighborhood to a neighborhood of 0 (because τ goes to 0 and the conjugate goes to ∞), and then we see that U becomes a sector of a circle. We then make that into a full circle is to use ρ, which raises $\delta(U)$ to some power. (This is the identity except at elliptic points, which have h_T greater than 1.) Notably, we make the neighborhoods small enough that the U_τ avoid other elliptic points.

We need to figure out how to define our charts for $\mathbb{H}$ as well. We'll do the same strategy: starting with a straightening map, and then identifying things that are Γ-equivalent. If we look at this picture from ∞, we can consider the neighborhoods $\mathcal{N}_2 \cup \{\infty\}$. Under $\Gamma = X(1)$, this neighborhood is periodic every 1, and that is the only equivalence. In general, this is periodic every h, and thus we just need to use the map

$$\rho(\tau) = e^{2\pi i \tau / h}.$$

This is sort of like the q-series map earlier, where we sent our map into a Fourier series on the unit disk. So if we take an $s \in \mathbb{Q} \cup \{\infty\}$, we will send our point to ∞ first, and then we'll use the identification map ρ to fill out the whole disk. (We might use different h_s's for different points.)

Formally, we're defining

$$h_s = |SL_2(\mathbb{Z})_\infty / (\delta\{\pm 1\}\Gamma\delta^{-1})_\infty|.$$

Basically, we're extracting out the fact that $(\delta\Gamma\delta^{-1})_\infty$ is generated by $\begin{bmatrix} 1 & h_s \\ 0 & 1 \end{bmatrix}$. It should be clear that h_s is finite – otherwise, we get an infinite index of $SL_2(\mathbb{Z})$.

Defining ψ to be the composition $\phi \circ \pi$, we know that ψ is not necessarily a bijection. However, we can show that $\pi(U)$ (which is our modular curve) is in bijection with V, because only the Γ-equivalent points are sent to the same point. π and ψ are both open continuous maps, so π is open and continuous as well, which shows that ϕ is indeed a homeomorphism.

To show holomorphicity, the idea is that "maps look more or less holomorphic." When we have one point in $\mathbb{H}$ and one in $\mathbb{Q} \cup \{\infty\}$, consider the map from V_1 to U_1, which sends q to $\delta_1^{-1} q^{1/h_1}$. After applying this map, we go from U_1

to U_2 by applying some γ, and then we go from U_2 to V_2 by applying the exponential map $e^{2\pi i\delta_2 x/h_2}$. **Everything here is holomorphic** except when we're taking the h_1th roots. But our domain never contains 0, because that means that $U_1 \cap U_2$ contains an elliptic point, and then $\pi(U_2)$ contains the elliptic point $\pi(\tau)$ (and we defined our neighborhoods so that the only elliptic points are possibly at the center), so we can always define q^{1/h_1} in a consistent way.

For the case where we have two cusps, we know that $\delta_2\gamma\delta_1^{-1}$ is a translation if our neighborhoods intersect, and then we just get another chain of holomorphic maps – this case isn't too bad.

And now we're done – we've showed all of the required conditions for X being a compact Riemann surface.

We can now state a first version of the Modularity Theorem:

Theorem 184

Let E be a complex elliptic curve with rational j-invariant. Then there is a positive integer N such that there is a surjective holomorphic function $X_0(N) \to E$, known as a **modular parameterization**.

15 April 9, 2020

Diamond and Shurman 3.1 – Christian Altamirano

We'll talk today about the **genus** of compact Riemann surfaces. There is a rigorous definition for genus, but intuitively, it is just the number of "holes" in our surface (so 0 for a sphere and 1 for a torus). Our goal will be to compute this number for a few Riemann surfaces.

Recall that the **modular curve** $X(\Gamma)$ is defined to be the set of orbits $\{\Gamma\tau : \tau \in \mathbb{H}^* = \mathbb{H} \cup \mathbb{Q} \cup \{\infty\}\}$. We showed earlier that $X(\Gamma)$ is a Riemann surface, and whenever we have a nonconstant holomorphic map $f : X \to Y$ between compact Riemann surfaces, f is surjective, and $f^{-1}(y)$ is discrete (so finite) for any $y \in Y$.

It turns out that there is a well-defined **degree** $d \in \mathbb{Z}^+$ for our function f, so that $|f^{-1}(y)| = d$ for all but finitely many points $y \in Y$. We'll be proving something a bit more general:

Definition 185

For any $x \in X$, let $e_x \in \mathbb{Z}^+$ be the **ramification degree** of f at x. In other words, e_x is the multiplicity with which f takes 0 to 0 as a local map, meaning $g(x) = x^{e_x}h(x)$ and $h(x) \neq 0$.

Recall the following theorem from complex analysis:

Theorem 186 (Local Mapping Theorem)

Let f be a holomorphic function, and suppose that $f(z) - w_0$ has a zero of order n (that is, with ramification degree $e_x = n$). Then points w near w_0 will have n distinct roots for the solution $f(z) = w$ near z.

Lemma 187

There exists d such that

$$\sum_{x \in f^{-1}(y)} e_x = d$$

for any $y \in Y$.

Proof. Let $\mathcal{E}$ be the set of **exceptional points** – that is, the points with ramification degree more than 1. This set is finite, because all such points are roots of f', which has finitely many zeros. Therefore, $X' = X \setminus \mathcal{E}$ and $Y' = Y \setminus f(\mathcal{E})$ are both still connected. Now fix $y \in Y'$; we know that for each $x \in f^{-1}(y)$, we have a neighborhood U_x such that f is locally bijective on U_x (because x has ramification degree 1); shrink these neighborhoods so that they are disjoint from each other, and such that they map to the same neighobrhood $V \in Y'$ containing y. (This is okay because there are finitely many images in the preimage.) Now we can define a function $y \to |f^{-1}(y)|$ on V; this is a continuous function and it is integer-valued, so it must be constant.

Let that constant be d: we now know that $\sum_{x \in f^{-1}(y)} e_x = d$ for all points $y \in Y'$ (everything except the exceptional points). To extend this to Y, note that whenever $y = f(x)$ and x is an exceptional point, we can find a neighborhood $N(y)$ of y such that every point in that neighborhood has $\sum_{x \in f^{-1}(y)} e_x = d$. (Basically, we're replacing e_x points of multiplicity 1 with 1 point of multiplicity e_x). Thus we have the result that we want. □

Definition 188

Define the **degree** of $f : X \to Y$ be the unique $d \in \mathbb{Z}^+$ such that

$$\sum_{x \in f^{-1}(y)} e_x = d$$

for all $y \in Y$.

Theorem 189 (Riemann-Hurwitz)

Let g_X, g_Y be the genera of two compact Riemann surfaces X and Y. Then

$$2g_X - 2 = d(2g_Y - 2) + \sum_{x \in X} (e_x - 1).$$

Proof sketch. The idea is to triangulate Y – we get E_Y edges and F_Y faces, and we know that $2 - 2g = F - E + V$ for any surface. Then lifting under f^{-1} yields a triangulation of X with dE_Y edges and dF_Y faces, but we lose $\sum_x (e_x - 1)$ vertices due to ramification. □

So now we can return to modular curves: suppose that $\Gamma_1 \subset \Gamma_2$ are congruence subgroups of $SL_2(\mathbb{Z})$. There is a natural projection $f : X(\Gamma_1) \to X(\Gamma_2)$, sending the orbits $\Gamma_1 \tau \to \Gamma_2 \tau$. We can think of this as a nonconstant holomorphic map between Riemann surfaces, and we can therefore calculate the degree:

Proposition 190

We have

$$\deg(f) = [\{\pm I\}\Gamma_2 : \{\pm I\}\Gamma_1] = \begin{cases} [\Gamma_2 : \Gamma_1]/2 & -I \in \Gamma_2, -I \not\in \Gamma_1 \\ [\Gamma_2 : \Gamma_1] & \text{otherwise.} \end{cases}$$

Proof. Partition $\{\pm I\}\Gamma_2$ into a coset partition $\bigcup_j \{\pm I\}\Gamma_1 y_j$. Pick a point $\Gamma_2 \tau$ (an orbit in $X(\Gamma_2)$) such that this is not the image of a point that ramifies; we'll show that $f^{-1}(\Gamma_2 \tau) = \{\Gamma_1 \gamma_j \tau\}$. (There are only finitely many points that ramify.)

We verify both inclusions: $\Gamma_1 \gamma_j \tau \in f^{-1}(\Gamma_2 \tau)$, because we'll take $f(\Gamma_1 \gamma_j \tau)$ to $\Gamma_2 \gamma_j \tau = \Gamma_2 \tau$, and for any $\Gamma_1 \tau' \in f^{-1}(\Gamma_2 \tau)$, we know that $f(\Gamma_1 \tau') = \Gamma_2(\tau') = \Gamma_2 \tau$, and then we can find γ such that $\tau' = \gamma \tau$, meaning that $\Gamma_1 \tau' = \Gamma_1 \gamma_j \tau$.

There is no ramification here, so $e_x = 1$ for every point in $f^{-1}(\Gamma_2\tau)$, meaning that $\deg f$ is the number of cosets. □

Here, recall that we multiply by $\{\pm I\}$ in the congruence subgroups, because $-I$ fixes any point and the action of $SL_2(\mathbb{Z})$ factors through $\pm I$.

Earlier on in the class, we discussed the local structure on Riemann surfaces with the straightening maps δ and wrapping maps $\rho_1(z) = z^{h_1}, \rho_2(z) = z^{h_2}$. For a subset U of $\mathbb{H}$, define $\rho_1 \circ \delta(U) = V_1$ and $\rho_2 \circ \delta(U) = V_2$. If we denote $\Gamma_{j,\tau}$ to be the isotropy subgroup of τ in $j = 1, 2$, and we proved last time that $h_j = |\{\pm I\}\Gamma_{j,\tau}|/2 \in \{1, 2, 3\}$ (because elliptic points have order either 2 or 3). Since $\frac{h_2}{h_1}$ should be integral, because the local map from V_1 to V_2 is the holomorphic map q^{h_2/h_1}, there are very few cases: we must either have $h_1 = 1$ or $h_1 = h_2$.

Proposition 191

The ramification degree for $\tau \in \mathbb{H}$ is h_2 when τ is an elliptic point for Γ_2 but not for Γ_1 and 1 otherwise. This is also the size of the quotient subgroup $[\{\pm I\}\Gamma_{2,\tau} : \{\pm I\}\Gamma_{1,\tau}]$.

Similarly, we can define a ramification degree for cusps. Earlier, we showed that if U is a neighborhood of $s \in \mathbb{Q}\cup\{\infty\}$, our maps $\rho_1(z), \rho_2(z)$ are $e^{2\pi i z/h_1}$ and $e^{2\pi i z/h_2}$, respectively, so this time the local map looks like $q \to q^{h_1/h_2}$, where $h_j = [SL_2(\mathbb{Z})_\infty : \{\pm I\}\Gamma_{j,s}]$ are the widths.

Proposition 192

The ramification degree for an $s \in \mathbb{Q} \cup \{\infty\}$ is $\frac{h_1}{h_2}$. This is also the size of the quotient subgroup $[\{\pm I\}\Gamma_{2,s} : \{\pm I\}\Gamma_{1,s}]$.

Proposition 193

Let Γ_1 be normal in Γ_2. Then all points in $X(\Gamma_1)$ with the same image in $X(\Gamma_2)$ have the same ramification degree.

Proof. Suppose $\Gamma_1\tau_1$ and $\Gamma_1\tau_2$ are points in $X(\Gamma_1)$ that map to the same point in $X(\Gamma_2)$. We know that $\Gamma_2\tau_1 = \Gamma_2\tau_2$, so we can find γ such that $\gamma\tau_1 = \tau_2$. And the idea from here is that conjugation by γ does not change the period. □

This now allows us to compute the genus: recall that for a group action G on X, Gx is the orbit of $x \in X$. We'll consider the case $\Gamma_1 = \Gamma$ and $\Gamma_2 = SL_2(\mathbb{Z})$.

Definition 194 (Local definitions)

Let $y_2 = SL_2(\mathbb{Z})i$, $y_3 = SL_2(\mathbb{Z})e^{2\pi i/3}$, and $y_\infty = SL_2(\mathbb{Z})\infty$. These are an elliptic point of period 2, elliptic point of period 3, and cusp of $X(1)$, respectively. Let $\varepsilon_2, \varepsilon_3$ be the number of elliptic points of Γ in $f^{-1}(y_2)$ and $f^{-1}(y_3)$, and let ε_∞ be the number of cusps of $X(\Gamma)$.

But any elliptic point $\Gamma\tau$ of period 2 has an image that is also an elliptic point of period 2. (If $h_1 > 1$, then $h_2 = h_1 > 1$). The only elliptic point of period 2 in X is y_2, so we indeed have $\Gamma\tau = y_2$. The same argument works for ε_3, so ε_2 and ε_3 account for all of the elliptic points.

So now for $h = 2, 3$, we know that the ramification degree of the points in $f^{-1}(y_h)$ is h for all elliptic points and 1 otherwise. Thus, the degree of f is

$$d = \sum_{x \in f^{-1}(y_h)} e_x = h \cdot (|f^{-1}(y_h)| - \varepsilon_h) + 1 \cdot \varepsilon_h.$$

Similarly, we can find ε_∞ by noting that

$$d = \sum_{x \in f^{-1}(y_\infty)} e_x \implies \sum_{x \in f^{-1}(y_\infty)} (e_x - 1) = d - \varepsilon_\infty,$$

and then we can apply Riemann-Hurwitz:

Theorem 195

The genus of $X(\Gamma)$ is

$$g = 1 + \frac{d}{12} - \frac{\varepsilon}{2} - \frac{\varepsilon}{3} - \frac{\varepsilon}{\infty},$$

where d is the degree of the natural projection $f : X(\Gamma) \to X(1)$.

(Here, we've used the fact that the fundamental domain D has genus 0 – it's basically a triangle.)

Theorem 196

Let p be a prime, and let $k = p + 1$. The genus of the congruence subgroup $X_0(p) = X(\Gamma_0(p))$ is

$$g = \begin{cases} \left\lfloor \frac{k}{12} \right\rfloor - 1 & k \equiv 2 \text{ mod } 12 \\ \left\lfloor \frac{k}{12} \right\rfloor & \text{otherwise.} \end{cases}$$

k will show up later as the weight of some modular forms – this should remind us of computing the dimension of cusp forms of some weight k.

Proof. We'll need a few results: let $\alpha_j = \begin{bmatrix} 1 & 0 \\ j & 1 \end{bmatrix}$ for $j \in [0, p-1]$, and let $\alpha_\infty = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}$. The point of these α_js is that we can split $SL_2(\mathbb{Z})$ into a disjoint union $\bigcup_j \Gamma_0(p)\alpha_j$.

We'll skip the proofs of a few computational results:

Lemma 197

$X_0(p)$ has exactly two cusps at 0 and ∞.

Lemma 198

For any $0 \le j < p$, we have $\gamma\alpha_j(i) = \alpha_j(i)$ for some $\gamma \in \Gamma_0(p)$ of order 4 if and only if $j^2 + 1 \equiv 0 \text{ mod } p$.

The purpose of these results is that the number of elliptic points of period 2 is the number of solutions to $x^2 + 1 \equiv 0 \text{ mod } p$. This is because we have a nontrivial martix γ that fixes $\alpha_j(i)$, so $\alpha_j(i)$ is an elliptic point, and because it is in the orbit of i, it must have period 2. And we know how to compute this number mod 4 using some results in algebra (it's 2 for $p \equiv 1$ mod 4, 0 for $p \equiv 3$ mod 4, and 1 for $p = 2$).

Lemma 199

For any $0 \le j < p$, we have $\gamma\alpha_j(e^{2\pi i/3}) = \alpha_j(e^{2\pi i/3})$ for some $\gamma \in \Gamma_0(p)$ of order 6 if and only if $j^2 - j + 1 \equiv 0 \text{ mod } p$.

The logic is the same here, and we also know how to compute this number mod 3. Putting all of this together means that we can do casework on the residue of p mod 12, to give the desired result. □

Example 200

Setting $p = 13$, this is the smallest prime where all four possible elliptic points occur.

We can also consider some coset representatives of $SL_2(\mathbb{Z})/\Gamma_0(13)$:

$$\beta_j = \begin{bmatrix} 1 & 0 \\ j & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \beta_\infty = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \alpha_\infty \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},$$

and this partitions the fundamental domain D of $X_0(13)$ into 14 regions. There are 13 points in $SL_2(\mathbb{Z})i$, and we can actually associate these with the $\beta_j(i)$: it turns out that $jj' + 1 \equiv 0 \mod 13$ if and only if $\gamma\beta_j(i) = \beta_{j'}(i)$ for some γ of degree 4.

We also yield some facts about elliptic points: for example, because $5 \cdot 5 + 1 \equiv 0 \mod 13$, this gives So this allows us to identify points of the boundary arc with each other, as well as understand the orientation of the fundamental domain.

Modular forms and representations of real groups – Professor Kim

We'll talk about how modular forms induce representations for real groups (like $SL_n(\mathbb{R}), SO_n(\mathbb{R})$, and so on).

Definition 201

Let G be a group (in general, we're interesting in locally compact groups such as $SL_2(\mathbb{R})$). A **representation** of G on a vector space V_π is a continuous homomorphism $\pi : G \to \text{Aut}(V_\pi)$.

Definition 202

Let G act on a topological space X, and let $V_\pi = C_C^\infty(x)$ (the space of compactly supported smooth functions). Define the **right regular representation** R_x via the action

$$(R_x(g)f)(x) = f(xg)$$

when $X = G$.

One of the questions to study is the space of L^2 (square-integrable) functions $L^2(G)$ or $L^2(\Gamma\backslash G)$ for some discrete subgroup. The point is that **modular forms (as well as Maass forms) generate representations of $SL_2(\mathbb{R})$, which induces an automorphic representation.**

To understand what's going on, let's fix a group $G = SL_2(\mathbb{R})$ and let $K = SO(2)$, $\Gamma = SL_2(\mathbb{Z})$. We know that G/K is homeomorphic to $\mathbb{H}$, because K can be characterized as the stabilizer of i in G, so we send g to gi. Now recall that a modular form is a holomorphic function with the automorphic condition

$$f(g, z) = j(g, z)^k f(z),$$

where $j = cz + d$. This identity can be notationally rewritten as $(f|_k g)(z) = f(z)$ – this is the $[\gamma]_k$ operator from earlier in the class – and now we can define a function

$$\phi_f(g) = f|_k(g)(i) = j(g, z)^{-k} f(gi).$$

This function ϕ_f inherits some properties from f: it is smooth (because f is nice), it defines a function on the quotient space $C^\infty(p\backslash SL_2(\mathbb{R})$, and the right action of the rotation matrix $r_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \in K$ pops out as a "character" in k:

$$\phi_f(gr_\theta) = e^{rk\theta}\phi_f(g).$$

This means that ϕ_f is a one-dimensional representation of K.

Also, since f is holomorphic at ∞, there is moderate growth of this function ϕ_f at $i\infty$. (In addition, if f is cuspoidal, then ϕ_f is rapidly decreasing.) Finally, because f is holomorphic, $\frac{\partial f}{\partial \overline{z}} = 0$, which means there exists a differential operator F such that $F \cdot \phi_f = 0$.

In general, the above properties define automorphic forms of $SL_2(\mathbb{R})$, and more generally any **real reductive group**.

Definition 203

Let G be a connected real reductive group. and let Γ be an arithmetic subgroup of G. A function ϕ on $\Gamma\backslash G$ is an **automorphic form** if ϕ is smooth, of moderate growth, is K-finite (the vector space $K\phi$ is finite-dimensional), plus another property related to Maass forms and the Lie algebra.

We can let $A(G, \Gamma)$ be the space of automorphic forms on G with respect to Γ. This is G-stable with respect to the right-regular representation, and we can use this to study $L^2(G)$ because the cuspoidal automorphic forms are contained in $L^2(G)$: $A_0(G, \Gamma) \subset L^2(G)$. If we go back to the $SL_2(\mathbb{Z})$ example, take a function $f \in \mathcal{M}_k(\Gamma)$, and consider

$$V_{\phi_f} = \langle R(g)\phi_f | g \in G\rangle.$$

This gives us a **holomorphic discrete series** π_k of minimal weight k, and we can make the connection here:

$$M_k(\Gamma) \cong \mathrm{Hom}_{(g,K)}(\pi_k A(G, \Gamma)).$$

16 April 14, 2020

Diamond and Shurman 3.2 – Zack Chroman

We'll be talking today about **automorphic forms**. Recall from earlier in the class the **weight-k operator**

$$(f[\alpha]_k)(\tau) = j(\alpha, \tau)^{-k} f(\alpha(\tau)),$$

where $j(\alpha, \tau) = c\tau + d$ and $\alpha \in SL_2(\mathbb{Z})$. In Serre, we defined a modular function in order to define a modular form, and we went from that to restricting to specific congruence subgroups. We'll work backwards now:

Definition 204

An **automorphic form** is a function $f : \mathbb{H} \to \mathbb{C}$ that is meromorphic, weight-k invariant under the congruence subgroup Γ, and meromorphic at infinity for all $\alpha \in SL_2(\mathbb{Z})$. Denote the space of automorphic forms at weight k as $\mathcal{A}_k(\Gamma)$.

When we look at the case of $k = 0$. $\mathcal{A}_0(\Gamma) \cong \mathbb{C}(X(\Gamma))$ corresponds to the space of meromorphic functions on $X(\Gamma)$. This is because our functions need to be invariant under Γ, so we can define on a reduced domain $X(\Gamma)$.

Theorem 205

We have that $\mathcal{A}_0(SL_2(\mathbb{Z}) = \mathbb{C}(j)$, where $j = \frac{1728g_2^3}{\Delta}$.

Here, $\mathbb{C}(j)$ is the **space of rational functions in j.**

Proof. Clearly $\mathbb{C}(j)$ is contained in $\mathcal{A}_0(SL_2(\mathbb{Z})$, because j is weight zero. To go backwards, we'll need to do more work. We know that the Laurent series is of the form

$$j(\tau) = \frac{1}{q} + \sum_{n=0}^{\infty} a_n q^n,$$

so the only pole of j is at infinity, because that's the only point where $\Delta = 0$. So j is holomorphic on the space Y_1 (which is X_1 without the point at infinity), which means $j : Y(1) \cong \mathbb{C}$ is a homeomorphism. In general, if we have a meromorphic function f which has zeros z_i and poles p_j, we can define

$$g(\tau) = \frac{\prod_{i=1}^{a}(j(\tau) - j(z_i))}{\prod_{i=1}^{b}(j(\tau) - j(p_i))}.$$

This function has the same zeros and poles as f, so $\frac{f}{g}$ has no poles and zeros at any finite point. And the number of poles and zeros need to add up to the same number on this compact Riemann surface, including multiplicity, so the ratio must have no zeros or poles even at infinity, which means that $f = cg$ for some constant c. Since g is a rational function of τ, this means f is also a rational function of τ. □

Later on, we'll be able to compute $\mathcal{A}_0(\Gamma)$ for other congruence subgroups, too.

Our next goal is to extend this result – when $k \neq 0$, we can't make our automorphic forms into a map on $X(\Gamma)$, because even if we have two points that are equivalent under Γ, they'll have some conjugate value under the action. But we should be able to define the **order of vanishing** consistently, and what's morally going on is that two points under the action of $\gamma \in \Gamma$ will keep the order the same because the factor of automorphy $cz + d$ has no zeros or poles on $\mathbb{H}$.

Definition 206

If $\pi(\tau)$ (the projection from $\mathbb{C}$ to $X(\Gamma)$) is not a cusp, define the **order of vanishing** at $\pi(\tau)$ as

$$\nu_{\pi(\tau)}(f) = \frac{\nu_\tau(f)}{h},$$

where $h = |\pm I\Gamma_\tau / \pm I|$ is the period of the elliptic point τ.

The idea is that the coordinate charts are no longer injective – they're h-to-one. So this order of vanishing is sometimes not an integer, if we're looking at elliptic points where we have half- or third-integers.

But the case where $\pi(\tau)$ is a cusp, the order of f at τ is more complicated. We'll consider $\tau = \infty$ first: recall that we needed the **local coordinate** $q = e^{2\pi i\tau/h}$, and then we wrote f as a power series in q, defining the order to be the lowest degree term. Here, h is the **width**: it's the minimal h such that one of $\pm \begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix}$ are in Γ.

If f is h-periodic, we can write down a power series in q, and everything works out. But it's possible that $-\begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix}$ is in Γ, but $\begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix}$ is not. Then $f(\tau + h) = -f(\tau)$, meaning our function is $2h$-periodic instead of h-periodic.

So we'll define $h' = 2h$, so that we're writing our power series as

$$f(\tau) = \sum_{n=m}^{\infty} a_n q_{h'}^n,$$

and then we'll define the **order at infinity** $\nu_\infty(f) = \frac{h}{2}$.

And to get the complete definition at cusps, we just need to conjugate from ∞ to s:

Definition 207

Let s be a cusp. Then the **order of vanishing** of $\pi(s)$ is

$$\nu_{\pi(s)}(f) = \begin{cases} \nu_s(f)/2 & k \text{ odd, } (\alpha^{-1}\Gamma\alpha)_\infty = \left\langle -\begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix} \right\rangle \\ \nu_s(f) & \text{otherwise,} \end{cases}$$

where α is the element of $SL_2(\mathbb{Z})$ such that $\alpha\infty = s$.

In almost all cases, f will be h-periodic, and we won't have the extra case. We'll call the extra case an **irregular cusp**, and we'll see later on that the only case this happens for our familiar congruence subgroups is $s = \frac{1}{2}$ for $\Gamma_1(4)$.

From here, we're done with the definitions, and we now have a nice result to characterize the space of cusp forms:

Theorem 208

Let $k, N \in \mathbb{N}$ such that $k(N+1) = 24$, and let $\mathcal{S}_k$ be the space of all cusp forms of weight k. Define $\phi_k(\tau) = \eta(\tau)^k\eta(N\tau)^k$, such that $\eta(\tau) = e^{2\pi i/24}\prod_{n=1}^{\infty}(1-q^n)$.

- If $\mathcal{S}_k(\Gamma_1(N))$ is nonzero, it is equal to $\mathbb{C}\phi_k$.
- In addition, if $S_k(\Gamma_0)$ is nonzero, then $\mathcal{S}_k(\Gamma_0(N)) = S_k(\Gamma_1(N)) = \mathbb{C}\phi_k$.

Here, N is either 1 or a prime number, so the second point follows from the first because $\mathcal{S}_k(\Gamma_0(N)) \subseteq \mathcal{S}_k(\Gamma_1(N))$. An important case is $N = 1, k = 12$: this tells us that all cusp forms of weight 12 over $SL_2(\mathbb{Z})$ are multiples of Δ.

Proof. We can first define

$$g = \phi_k^{N+1} = (2\pi)^{-24}\Delta(\tau)\Delta(24\tau),$$

and some more substitution yields

$$= q^{N+1}\prod_{n=1}^{\infty}(1-q^n)^{24}(1-q^{Nn})^{24}.$$

Now $\Delta(N\tau)$ is a cusp form on $\mathcal{S}_{12}(\Gamma_0(N))$, so substitution tells us that $g \in \mathcal{S}_{24}(\Gamma_0(N)) \subseteq \mathcal{S}_{24}(\Gamma_1(N))$.

Lemma 209

Our function g has a zero of order $N+1$ at every cusp $\pi(s)$ (here, π is the projection map onto $X(\Gamma_1(N))$).

Proof of lemma. We'll need to blackbox a few results: it turns out that all cusps are regular for $N \neq 4$. Also, the cusps of $X(\Gamma_0(N))$ are $\pi_0(\infty)$ and $\pi_0(0)$, and the widths here are 1 and N respectively. Finally, we'll need the projection map $X(\Gamma_1(N)) \to X(\Gamma_0(N))$ to be unramified at the cusps.

The second assumption is the most important here, because it gives us a handle on the cusps on the larger space $X(\Gamma_1(N))$. For any cusp $s \in X(\Gamma_1(N))$, we know that $\pi(s)$ lies over either $\pi_0(\infty)$ or $\pi_0(0)$.

In the first case, where $\pi(s)$ is a cusp over $\pi_0(\infty)$, we can write $s = \alpha\infty$ for some matrix $\alpha \in \Gamma_0(N)$. Since g is $\Gamma_1(N)$-invariant, it is invariant under the weight-24 action of α, which means that

$$\nu_{\pi(s)}(g) = \nu_s(g) = \nu_\infty(g) = N + 1,$$

where the last equality comes from the direct expansion of g.

In the second case, $\pi(s)$ is a cusp over $\pi_0(0)$, so the weight-24 operator of $S = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ on g looks like

$$g[S]_{24} = N^{-12} q_N^{N+1} \prod_{n=1}^{\infty} (1 - q_N^n)^{24} (1 - q_N^{Nn})^{24},$$

where $q_N = e^{2\pi i \tau / N}$. Since S sends ∞ to zero, the order of $\pi(S)$ at g is the same as the order of $\pi(0)$ at g, which is also $N + 1$. $\square$

One important point is that we needed to write things in terms of q_N instead of our usual q, because the width of $\pi_0(0)$ is N. So now we can finish the proof: if $f \in \mathcal{S}_k(\Gamma_1(N))$, we can consider the automorphic form $\frac{f^{N+1}}{g}$. This is holomorphic on $Y(\Gamma_1(N))$, because g has no zeros (it's the product of Δ functions), and because of our lemma, this is also holomorphic on the cusps. This is because the numerator f^{N+1} has a zero at every cusp, so the order is at least as large in the numerator as the denominator. But then this is a holomorphic function defined on a compact set $X(\Gamma_1(N))$, so we actually have that $\frac{f^{N+1}}{g} = c$ for a constant c, which implies that $f = c'\phi_k \in \mathbb{C}\phi_k$ (ϕ_k is one of the $(N+1)$th roots of g, and we need to pick the same $(N+1)$th root by continuity throughout). $\square$

We haven't shown that ϕ_k is actually a cusp form yet – we do know this is true for $(N, k) = (12, 1)$, because then $\mathcal{S}_k(SL_2(\mathbb{Z})) = \mathbb{C}(\Delta)$. On the other hand, whenever k is odd, $\mathcal{S}_k(\Gamma_0(N)) = 0$, because $-I$ is in $\Gamma_0(N)$. But it does turn out that (using the dimension formulas that we'll go over in the next few classes) we do have

$$\mathcal{S}_k(\Gamma_0(N)) = \mathcal{S}_k(\Gamma_1(N)) = \mathbb{C}\phi_k :$$

we have a space of cusp forms of dimension 1.

Diamond and Shurman 3.3 – Michael Tang

The topic of this section is **meromorphic differentials** – these will be helpful in talking about modular forms for the rest of the chapter. Today, we'll go over some motivation, define local and then global differentials, and then we'll go from differentials to automorphic forms. On Thursday, we'll finish the lecture by going in the opposite direction, giving us an isomorphism between automorphic forms and differentials, and we can use this to talk about dimension of modular forms.

We'll start with the transformation rule: for any function $f \in \mathcal{A}_{2n}(\Gamma)$, we have

$$f(\gamma(\tau)) = j(\gamma, \tau)^{2n} f(\tau) \quad \forall \gamma \in \Gamma.$$

We can notice that $\frac{d\gamma(\tau)}{d\tau} = j(\gamma, \tau)^{-2}$ by computing with the quotient rule, so we can now say that

$$f(\gamma(\tau)) = \left(\frac{d\gamma(\tau)}{d\tau} \right)^{-n} f(\tau),$$

and now we can symbolically treat the derivative as a fraction and rewrite as

$$f(\gamma(\tau)) = (d\gamma(\tau))^n = f(\tau)(d\tau)^n.$$

This doesn't have any real meaning yet, but we'll make it more formal later. But this is nice, because $f(z)(dz)^n$ is invariant under the action of Γ – we'll make the definitions make sense now, so that we can exploit the useful properties here.

We'll start by looking at open sets on $\mathbb{C}$:

Definition 210

Let $V \subseteq \mathbb{C}$ be an open set, and let $n \in \mathbb{N}$. The **meromorphic differentials on V of degree n** are

$$\Omega^{\otimes n}(V) = \{f(q)(dq)^n : f \text{ meromorphic on } V\}.$$

We can set up a vector space structure over $\mathbb{C}$, because f and g being meromorphic means fg, cf are both meromorphic as well. The dq here is just a symbol – everything is really just determined by f and n.

In order to study this on Riemann surfaces, we'll need to talk about a mapping between local differentials: if we have two open sets V_1, V_2 and a holomorphic map ϕ between them, we'll induce a **pullback map** ϕ^* in the other direction between the differentials, defined by

$$\phi^*(f(q_2)(dq_2)^n) = [f(\phi(q_1))(\phi'(q_1))^n](dq_1)^n.$$

This is basically a u-substitution. We can confirm that whenever ϕ is holomorphic and f is meromorphic, $(f \circ \phi)(\phi')^n$ is also meromorphic.

Example 211

If we take the inclusion $\iota : V_1 \subseteq V_2$, then ι^* is just a restriction:

$$\iota^*(f(q)(dq)^n) = [f(\iota(q))1^n](dq)^n = f|_{V_1}(q)(dq)^n.$$

We can note a few properties of this pullback map:

- ϕ^* is a linear map from $\Omega^{\otimes n}(V_2)$ to $\Omega^{\otimes n}(V_1)$.
- ϕ^* is a **contravariant operator**: $(\phi_2 \circ \phi_1)^* = \phi_1^* \circ \phi_2^*$.

Both of these results are just bookkeeping of our earlier definitions. We'll also state a few properties without proof:

- If ϕ is surjective, then ϕ^* is injective. Similarly, if ϕ is a bijection, so is ϕ^*, and we have the identity $(\phi^*)^{-1} = (\phi^{-1})^*$.

We now want to define differentials on the entire Riemann surface X. Recall that when we defined the Riemann surface, we used a collection of coordinate charts $(\phi_j)_{j \in J}$, each of which maps neighborhoods U_j of X (here the U_j must cover X) to open sets in $\mathbb{C}$, and one condition is that they must be compatible:

$$\phi_{k,j} = \phi_k \circ \phi_j^{-1}$$

must be a holomorphic map from $V_j \to V_k$ for all $j, k \in J$. So similarly, a global differential should be a collection of local differentials which are compatible.

Definition 212

Let X be a Riemann surface with charts $\phi_j : U_j \to V_j$. A **meromorphic differential on X of degree n** is a collection of $(\omega_j)_{j \in J}$, where $\omega_j \in \Omega^{\otimes n}(V_j)$, such that for all $j, k \in J$,

$$\phi_{k,j}^* \left(\omega_k |_{\phi_k(U_j \cap U_k)} \right) = \omega_j |_{\phi_k(U_j \cap U_k)}$$

(the pullback of the transition map does send ω_k to ω_j whenever both are defined). We denote the set of differentials by $\Omega^{\otimes n}(X)$.

Example 213

Consider $X = \mathbb{C}/\Lambda$ to be a complex torus. Recall that the coordinate charts ϕ_j are just the projection map $\pi : \mathbb{C} \to X$, and the inverses are just defined locally in the simple way. Then the transition maps are then just "corrections between fundamental domains:" we have $\phi_{k,j}(z) = z + \lambda$ for some $\lambda \in \Lambda$.

We can check the compatibility condition – because the inverse should have derivative 1, we can just say that $\omega_j = dz$ for all $j \in J$, and then indeed $\phi_{k,j}^*(\omega_j) = dz = \omega_k$, and the collection of ωs makes sense on X.

So now we can work towards automorphic forms: let Γ be a congruence subgroup of $SL_2(\mathbb{Z})$, and we'll spend the rest of this lecture on $X(\Gamma)$. Remember that neighborhoods of $X(\Gamma)$ look like $\pi(U_j)$, where π is a projection from $\mathbb{H}^*$ to $X(\Gamma)$ and U_j. Then we defined the coordinate charts indirectly via

$$\psi_j = \phi_j \circ \pi,$$

where $\psi_j = e^{2\pi i \delta(\tau)/h}$ is the identification map we've been disussing. So the local differentials are defined on V_j, and now we can define a **global pullback** of ω by pulling back the ω_js individually:

$$\pi^*(\omega)|_{U_j'} = \psi_j^*(\omega_j'),$$

where the primes mean that we're only working on $\mathbb{H}$. We just need to make sure these local pullbacks agree on the intersections of the U_js, in the same way that we needed to make sure the differentials were compatible. In other words, **we need to show that**

$$\psi_j^* \left(\omega_j |_{V_{j,k}} \right) = \psi_k^* = \psi_k^* \left(\omega_k |_{V_{j,k}} \right),$$

where $V_{j,k} = \psi_j(U_j' \cap U_k')$ and $V_{k,j} = \psi_k(U_j' \cap U_k')$.

Proof. We know that $\psi_j = \phi_j \circ \pi$ and $\psi_k = \phi_k \circ \pi$, so solving for π yields

$$\psi_j \circ \phi_j^{-1} = \pi = \psi_k \circ \phi_k^{-1}.$$

This means that

$$\phi_{kj} = \phi_k \circ \phi_j^{-1} = \psi_k \circ \psi_j^{-1},$$

meaning that we can write $\psi_k = \phi_{kj} \circ \psi_j$. This is helpful, because the compatibility condition for the ωs has to do with the pullback of the transition map. So taking the pullback of both sides and using contravariance yields

$$\psi_k^* = \psi_j^* \circ \phi_{kj}^* \implies \psi_k^*(\omega) = \psi_j^*(\phi_{kj}^*(\omega_k)) = \psi_j^*(\omega_j),$$

as desired. □

This shows that π^* is well-defined, because the local pullbacks take the same value whenever they intersect. And the U_j's cover $\mathbb{H}$ – we'll talk on Thursday about extending this to the cusps, too.

In summary, if we have a meromorphic differential $\omega \in \Omega^{\otimes n}(X(\Gamma))$ of degree n, we can construct a meromorphic differential $\pi^*(\omega)$ on $\mathbb{H}$, also of degree n. We can then define the function f such that

$$\pi^*(\omega) = f(\tau)(d\tau)^n.$$

Since $\pi^*(\omega)$ lives in $X(\Gamma)$, the pullback doesn't care about the individual elements of Γ, so it must be Γ-invariant: in other words, γ^* acts trivially on $\Omega^{\otimes n}(X(\Gamma))$.

And then

$$f(\tau)(d\tau)^n = \gamma^* \left(f(\tau)(d\tau)^n\right) = f(\gamma(\tau))(\gamma'(\tau))^n(d\tau)^n,$$

and plugging in the definition of the derivative in our case yields

$$f(\tau) = f(\gamma(\tau))j(\gamma,\tau)^{-2n},$$

and now we've gotten that f is **weakly modular** of weight $2n$, which is a (roundabout) way to motivate our definition of weakly modular!

Next time, we'll show that f is actually an automorphic form – we need to show meromorphicity when acting by the weight-k operators and at infinity. It turns out that it's easier to study differential forms than their corresponding automorphic forms, so this will help us get some more insight out of the group $\mathcal{A}_{2n}$.

17 April 16, 2020

Diamond and Shurman 3.3 continued – Michael Tang

We'll continue discussing meromorphic differentials today – as a recap, recall that we defined the local differentials $f(q)(dq)^n \in \Omega^{\otimes n}(V)$ on open sets $V \subseteq \mathbb{C}$, and then we defined **pullback maps** from $\Omega^{\otimes n}(V_2) \to \Omega^{\otimes n}(V_1)$ given a homomorphism between open sets $V_1 \to V_2$. We combined these local differentials into global differentials on Riemann surfaces, making sure that they were compatible through the pullback (transition) maps. Specifically, if we consider a congruence subgroup $\Gamma \subseteq SL_2(\mathbb{Z})$, we constructed a global pullback $f(\tau)(d\tau)^n$ on the modular curve $X(\Gamma)$. We noticed that this yields a weakly modular function of weight $2n$, and it turns out that $f[\alpha]_{2n}$ is actually meromorphic at ∞, so f is an element of $\mathcal{A}_{2n}(\Gamma)$. This is important, because it gives us a map between differentials and automorphic forms.

Today, we'll consider the inverse mapping as well: we'll start with $f \in \mathcal{A}_{2n}(\Gamma)$, and we'll construct the ω that pulls back to f. This will give us an isomorphism between the spaces of meromorphic forms and automorphic forms, and we'll see why this is important.

Remember that the mapping $\omega \to f$ was defined via the pullback map

$$\pi^*(\omega)|_{U_j'} = \psi_j^*(\omega_j') = f(\tau)(d\tau)^n.$$

Here, U_j' is an open subset of $\mathbb{H}$, π is the projection map, and ψ takes a neighborhood U_j to a neighborhood V_j. To show this was valid, we needed to show that the local pullbacks agreed: the compatibility condition we needed to satisfy was

$$\psi_k^*\omega_k) = (\psi_j^* \circ \phi_{kj}^*)(\omega_k) = \psi_j^*(\omega_j).$$

Let's show the converse is true as well:

Lemma 214

Suppose (ω_j) is a collection of differentials on the V_js, and let $\psi_J : U_j \to V_j$ be identification maps for $X(\Gamma)$. If $\psi_j^*(\omega_j') = \psi_k^*(\omega_k')$ on the intersections, then the ω_js are compatible, meaning they define a global differential.

Proof. Run the previous argument in reverse: each equality is an if and only if statement. □

This is helpful because we can start with an f, construct the individual ω_js, make sure they pull back to $f(\tau)(d\tau)^n$, and that will give us the ω.

For the inverse construction, let's start with an automorphic form $f \in \mathcal{A}_{2n}(\Gamma)$. We are doing things locally, so we can start with a neighborhood $U_j \subseteq \mathbb{H}^*$, we want to construct an ω_j so that

$$\psi_j^*(\omega_j) = (f(\tau)(d\tau)^n)|_{U_j}.$$

We'll want to do a "pushforward," but we can just try to take an inverse map ψ^{-1}. Unfortunately, ψ_j isn't invertible in general, so we'll need to look more carefully what's going on. Recall that when we set up the charts of Riemann surfaces, we wrote

$$\psi_j = \rho_j \circ \delta_j$$

where δ_j is a linear transformation (the action of the matrix $\begin{bmatrix} 1 & -\tau_j \\ 1 & -\overline{\tau_j} \end{bmatrix}$), while $\rho_j(z) = z^h$ for non-cusps and $\rho_j(z) = e^{2\pi i z/h}$ for cusps. So we can try to do our pushforward in two steps: construct an intermediate differential such that pulling back by δ_j gives the original differential ω_j, and then pulling back by ρ_j gives the intermediate differential.

We'll start with the linear transformation: luckily this is invertible, because the determinant is nonzero whenever $\tau_J \in \mathbb{H}$, so we just need to define

$$\lambda_j = \alpha^* \left((f(\tau)(d\tau)^n|_{U_j'} \right)$$

where $\alpha = \delta_j^{-1}$. And doing some computation with this,

$$\lambda_j = \alpha^* \left(f(\tau)(d\tau)^n|_{U_j'} \right) = f(\alpha(z))(\alpha'(z))^n(dz)^n,$$

and here $\alpha'(z) = \det \alpha j(\alpha, z)^{-2}$. (We can't assume $ad - bc = 1$ for the matrix like usual.) This yields

$$= f(\alpha(z)) \cdot (\det \alpha)^n j(\alpha, z)^{-2n}(dz)^n,$$

and we can still define (by analogy) that this $\lambda_j = (f[\alpha]_{2n})(z)(dz)^n$ – this is a **generalization of the weight-k operator**.

So now we need to find an ω_j such that $\rho_j^*(\omega_j) = \lambda_j$. If we take a small neighborhood U_j, we can try to find a formula for a local differential form that will pull back, at least in a small region. We'll just do the case where τ_j is not a cusp, so we have $\rho_j(z) = z^h$.

Lemma 215

The function $z^n(f[\alpha]_{2n})(z)$ is only dependent on z^h.

Remember that h is defined to be the size of the isotropy subgroup $\{\pm I\}(\delta_j \Gamma \delta_j^{-1})_{\tau_j}/\{\pm I\}$, and we showed that this is finite cyclic, generated by a rotation $r_h(z) = e^{2\pi i h}z$.

Proof. This follows from the Γ-invariance of $f(\tau)(d\tau)^n$, which means that $\alpha^*(f(\tau)(d\tau)^n))$ is $\delta_j\Gamma\delta_j^{-1}$-invariant:

$$(\delta_j\gamma\delta_j^{-1}) = (\delta_j^{-1})^* \circ \gamma^* \circ \delta_j^* = \alpha\gamma^*\alpha^{-1},$$

and applying this to α^* works out the way we want. And this means $r_h^*(\lambda_j) = \lambda_j$, so

$$(e^{2\pi i/h}z)^n(f[\alpha]_{2n})(e^{2\pi i/h}z) = z^n(f[\alpha]_{2n})(z),$$

which means that $z^n(f[\alpha]_{2n})(z)$ is invariant under a rotation by an hth root of unity, meaning it is only dependent on z^h. □

The point is that we can now define

$$g_j(z^h) = z^n(f[\alpha]_{2n})(z)$$

for some meromorphic function g_j. It turns out that

$$\omega_j = \frac{g_j(q)}{(hq)^n}(dq)^n$$

now works, because we can check that pulling back by ρ_j yields

$$\rho_j^*(\omega_j) = (f[\alpha]_{2n})(z)(dz)^n = \lambda_j,$$

because the $(hq)^n$ makes the derivatives work out. So now we've constructed the local differential ω_j!

Fact 216

For the cusp case, we can find a similar property: $(f[\alpha]_{2n})(z)$ turns out to only be a function of $e^{2\pi iz/h}$, and $\omega_j = \frac{g_j(q)}{(2\pi iq/h)^n}(dq)^n$.

So in all cases, we have a local differential ω_j, such that pulling back gives us our original $f(\tau)(d\tau)^n$. Since all the ω_js pull back to the same f, they are compatible, so we get a global differential on the Riemann surface.

Combining all of this together, we get the isomorphism we've been after:

Theorem 217

The space of automorphic forms is isomorphic to the space of meromorphic differentials: for any $n \in \mathbb{N}$ and congruence subgroup Γ,

$$\Omega^{\otimes n}(X(\Gamma)) \cong \mathcal{A}_{2n}(\Gamma)$$

as complex vector spaces.

We do have a bijection, and we just need to show that there is a linear map between them, but the pullback map is linear and we're not doing anything weird with multiplication or addition in ways that are not compatible.

This is important, because we want to compute the dimensions of the spaces $\mathcal{M}_k(\Gamma)$ and $\mathcal{S}_k(\Gamma)$ (modular forms and cusp forms, respectively). Here, **modular forms** are automorphic forms that are holomorphic, and **cusp forms** are such forms that vanish at ∞. Instead of looking at the dimensions directly, we can look at the images under the isomorphism, and compute the dimensions of the images in $\Omega^{\otimes k/2}(X(\Gamma))$. Riemann surfaces turn out to have more structure – we'll be able to write things in terms of the order of vanishing in a purely algebraic way. For example, being a modular form means that the order of vanishing of f at every point $\pi(\tau_j)$ and $\pi(s_j0$ is at least 0.

Diamond and Shurman 3.4 – David Wu

We'll follow up on the previous topic by talking about divisors and the Riemann-Roch formula. Throughout this lecture, we'll be talking about general Riemann surfaces X, though we really care about $X(\Gamma)$.

Definition 218

A **divisor** D on a compact Riemann surface X is a finite formal $\mathbb{Z}$-linear combination of points of X:

$$D = \sum_{x \in X} n_x x, \quad n_x \in \mathbb{Z}, n_x = 0 \text{ for all but finitely many } x.$$

We can look at the set of all divisors on X, denoted $\text{Div}(x)$, and this forms the free abelian group on the points of X: basically, we add two divisors by adding the coefficients. We'll define a (partial) ordering on $\text{Div}(x)$ as well: say that $D \geq D'$ if the coefficients satisfy $n_x \geq n'_x$ for all x.

Definition 219

The **degree** of a divisor is $\deg(D) = \sum_x n_x$ (since we assume that n_x is zero for all but finitely many x).

Notice that the degree map from $\text{Div}(x)$ to $\mathbb{Z}$ is a homomorphism of abelian groups. A natural question is to ask about the kernel of this map – one natural place to start is to look at the **function field** of meromorphic functions $\mathbb{C}(X)$. (Remember that being a meromorphic function f on X means that pre- and post-composing it with the embedding into $\mathbb{C}$ yields a meromorphic function on $\mathbb{C}$.) For any such meromorphic function f, we have the divisor

$$\text{div}(f) = \sum \nu_x(f) x,$$

where ν_f is the order of vanishing of f on x. Then we claim that the map

$$\text{div} : \mathbb{C}(X)^* \to \text{Div}(X)$$

is a homomorphism as well, because $\text{div}(f_1 f_2) = \text{div}(f_1) + \text{div}(f_2)$. (Just consider the Laurent expansion at each point: the powers should add when we multiply the functions.)

Proposition 220

The degree of any of these nonzero divisors $f \in \mathbb{C}(X)$ is zero.

Proof. A meromorphic function $f : X \to \mathbb{C}$ is a holomorphic function on $\hat{\mathbb{C}}$, where $\hat{\mathbb{C}}$ is $\mathbb{C}$ without a finite set of points. We know that (from Christian's lecture) the degree

$$d = \sum_{x \in f^{-1}(y)} \nu_x(\tilde{f})$$

is constant. But the number of zeros is equal to the number of poles if we take $y = 0, \infty$. □

As a more handwavy illustrative example, if we take X to be the Riemann sphere and we draw a closed contour, this is saying that we can apply the argument principle on both sides of the contour.

Definition 221

Let $\mathrm{Div}^\ell(X)$ be the group of divisors of nonzero meromorphic functions, and let $\mathrm{Div}^0(X)$ be the group of divisors of degree 0 (this is the kernel of the degree map). By the proposition above, $\mathrm{Div}^\ell(X)$ is a subgroup of $\mathrm{Div}^0(X)$.

There are divisors that don't come from meromorphic functions, but we have a handy theorem:

Theorem 222 (Abel)

Let g be the genus of X, and let Λ_g be a lattice that spans $\mathbb{C}^g$. Then

$$\mathrm{Div}^0(X)/\mathrm{Div}^\ell(X) \cong \mathbb{C}^g/\Lambda_g$$

(a g-holed torus).

For example, if we let $\Lambda_i = i\mathbb{Z} \oplus \mathbb{Z}$ and consider the elliptic curve $\mathbb{C}/\Lambda_i$, we can compute the divisor of $\frac{\wp'}{\wp}$, where $\wp$ is the Weierstrass function – we know the zeros and poles of this function, so this is easy to do directly.

Definition 223

For any divisor D on X, the **linear space** of D is

$$L(D) = \{f \in \mathbb{C}(X) : f = 0 \text{ or } \mathrm{div}(f) + D \geq 0\}.$$

This might look mysterious in general, but if we have $D = \mathrm{div}(\tilde{f})$ for some meromorphic function $\tilde{f}$, $L(D)$ consists of the functions f that "cancel out the poles," meaning $f\tilde{f}$ must be holomorphic. We can indeed check that

$$\nu_x(f_1 + f_2) \geq \min(\nu_x(f_1), \nu_x(f_2)),$$

so $L(D)$ is closed under addition. In fact, $L(D)$ is a finite-dimensional vector space: we'll denote its dimension to be $\ell(D)$. Take an $n \in \mathbb{N}$, and let $\omega \in \Omega^{\otimes n}(X)$ be a nonzero meromorphic differential on X. We have a local map at each point in $x \in X$ of the form

$$\omega_x = f_x(q)(dq)^n,$$

where f_x just means we're looking at the local map at the point x, and now if $\nu_x(\omega) = \nu_0(f_x)$ (because q is centered around 0), we define

$$\mathrm{div}(\omega) = \sum_x \nu_x(\omega)x.$$

Here, we'll have $\mathrm{div}(\omega_1\omega_2) = \mathrm{div}(\omega_1) + \mathrm{div}(\omega_2)$ for any $\omega_1 \in \Omega^{\otimes n}(X)$ and $\omega_2 \in \Omega^{\otimes m}(X)$.

Definition 224

A **canonical divisor on X** is a divisor of the form $\mathrm{div}(\lambda)$, where λ is a nonzero element of $\Omega^1(X)$.

Theorem 225 (Riemann-Roch)

Let X be a compact Riemann surface of genus g, and let $\mathrm{div}(\lambda)$ be any canonical divisor on X. For any $D \in \mathrm{Div}(X)$,

$$\ell(D) = \deg(D) - g + 1 + \ell(\mathrm{div}(\lambda) - D).$$

This is important because we have explicit information about canonical divisors. (We won't prove this.)

Corollary 226

Take previous notation. We have the following results:

1. $\ell(\text{div}(\lambda)) = g$.
2. $\deg(\text{div}(\lambda)) = 2g - 2$.
3. For any divisor D with $\deg(D) < 0$, we have $\ell(D) = 0$.
4. For any divisor D of degree larger than $2g - 2$, $\ell(D) = \deg(D) - g + 1$.

Proof. For (1), if f is nonconstant, it has a pole (for example, polynomials have a pole at infinity). This means $\text{div}(f) \geq 0$ cannot occur, which means that $L(0)$ only contains constant functions, meaning $\ell(0) = 1$. Taking $D = 0$ in Riemann-Roch, we find that $1 = -g + 1 + \ell(\text{div}(\lambda))$, so $\ell(\text{div}(\lambda))$ as desired.

For (2), we can set $D = \text{div}(\lambda)$, which yields $g = \deg(\text{div}(\lambda)) - g + 1 + 1$. Rearranging gives the desired result.

For (3), suppose that $\ell(D) > 0$, so there is a nontrivial element $f \in L(D)$. This means $\text{div}(f) \geq -D$, and taking degrees yields $\deg(D) \geq 0$, so we've proved the contrapositive.

Finally, (2) and (3) imply (4): we have $\ell(\text{div}(\lambda) - D) = 0$, and

$$\deg(\text{div}(\lambda) - D) = \deg(\text{div}(\lambda)) - \deg D < 0,$$

since the degree of D is larger than $2g - 2$. □

Recall that Zack's lecture showed us that a nonzero automorphic form $j' \in \mathcal{A}_2(\Gamma)$ exists for any congruence subgroup Γ. Let f be any nonzero automorphic form of weight 2: define

$$\lambda = \omega(f) \in \Omega^1(X(\Gamma))$$

to be the meromorphic differential for f. Then $\text{div}(\lambda)$ is a canonical divisor, and by (2) above, the degree must be $(2g - 2)$. So if we take any positive even integer k,

$$\lambda^{k/2} = \Omega^{\otimes k/2}(X(\Gamma))$$

(multiplying the meromorphic functions and the dqs), and because degree is a homomorphism, this gives us a degree of $k(g - 1)$ for the divisor of $\lambda^{k/2}$. We know the equality of vector spaces

$$\Omega^{\otimes k/2}(X(\Gamma)) = \mathbb{C}(X(\Gamma))\lambda^{k/2},$$

and for any meromorphic f, $\deg(\text{div}(f)) = 0$. Thus by additivity of the degree, **every nonzero differential** $\omega \in \Omega^{\otimes k/2}(X(\Gamma))$ has a degree of $k(g - 1)$.

So now if Γ is a congruence subgroup and g is the genus of $X(\Gamma)$, the space of holomorphic one-forms $\Omega^1_{\text{hol}}(X(\Gamma))$ is isomorphic to $L(\lambda)$, the linear space of λ. By (1) above, the dimension of the linear space $\ell(\lambda) = g$, so the dimension of $\mathcal{S}_2(\Gamma)$ is g. (We can show that the space of weight-2 cusp forms is isomorphic to the space of holomorphic differentials of degree 1.) In general, this argument will let us find general dimension formulas for $\mathcal{M}_k(\Gamma)$ and $\mathcal{S}_k(\Gamma)$.

In summary, we've defined these divisors in terms of the order of vanishing of the functions. When we proved the formulas for dimension on $SL_2(\mathbb{Z})$, we did an argument relating orders of vanishing at i, ρ, ∞, and so on, and we're generalizing here: we're taking the Riemann surface $X(\Gamma)$ and using order of vanishing data to relate that to the space of modular forms.

18 April 21, 2020

Diamond and Shurman 3.5 – Kaarel Haenni

The plan for today is to discuss dimension formulas for the spaces of modular forms and cusp forms of weight k: in this lecture, we'll combine results from the past few lectures to answer this question for even k.

We'll start with a quick review: recall that an **automorphic form** is a function $f : \mathbb{H} \to \hat{\mathbb{C}}$ which is like a modular form, but we only need to be meromorphic everywhere. If we take a nonzero $f \in \mathcal{A}_k(\Gamma)$, it may not make sense to talk about f as a function on $X(\Gamma)$ because it's not constant on orbits, but we can talk about the order of vanishing (which is constant). We define this order via $\nu_x(f) = \frac{\nu_\tau(f)}{h}$, where h is the period of τ. We want to talk about the formal sum

$$\text{div}(f) = \sum_{x \in X(\Gamma)} v_x(f)x,$$

which looks like the **divisor** from last lecture, but the small technical difficulty is that these v_x may be rational rather than just being integers. So we'll just fix this: we'll define the **rational-coefficient divisor space** $\text{Div}_{\mathbb{Q}}(X)$ to be the space of formal sums $\sum n_x x$, where n_x are rational and almost all zero. This still has a natural abelian group, a $\geq$ relation, and a degree function (adding all the coefficients, just like for the integer case).

Riemann-Roch does not extend directly to this $\mathbb{Q}$ case, but it will still be useful. Recall that a corollary of Riemann-Roch is that when we have a compact Riemann surface of genus g, the linear space $L(D)$ of a divisor $D \in \text{Div}^0(X)$ (which is the space of divisors of degree 0) satisfies

$$\ell(D) = \dim(L(D)) = \deg(D) - g + 1$$

whenever $\deg(D) > 2g - 2$. We also know that we have an isomorphism of complex vector spaces

$$\omega : A_k(\Gamma) \to \Omega^{\otimes k/2}(X(\Gamma)),$$

such that f is sent to ω_j and ω_j pulls back to $f(\tau)(d\tau)^{k/2}$. With this, we'll move on to new material: our goal is to reduce the dimension calculations to finding **dimensions of linear spaces**. We'll look at $k \geq 2$ **and even**, so that we know there exists a nonzero function $f \in \mathcal{A}_k(\Gamma)$. Let $\mathbb{C}(X(\Gamma))$ be the field of meromorphic functions on $X(\Gamma)$: recall that

$$\mathcal{A}_k(\Gamma) = \mathbb{C}(X(\Gamma))f$$

(the automorphic forms are equal to the meromorphic forms times a particular function). Using this description, we can describe the space of modular forms via

$$\mathcal{M}_k(\Gamma) = \{f_0 f \in A_k(\Gamma) \text{ holomorphic } \iff f_0 f = 0 \text{ or } \text{div}(f_0 f) \geq 0\},$$

which is isomorphic to the complex vector space

$$\{f_0 \in \mathbb{C}(X(\Gamma)) : f_0 = 0 \text{ or } \text{div}(f_0) + \text{div}(f) \geq 0\}$$

where we've used the fact that degree is a homomorphism. This is just the linear space of $\text{div}(f)$, and it seems that we might want to use Riemann-Roch. But $\text{div}(f)$ might have non-integer coefficients, so instead we'll just **take the floor**, $\lfloor \text{div}(f) \rfloor$, which is defined via

$$\left\lfloor \sum n_x x \right\rfloor = \sum \lfloor n_x \rfloor x.$$

Since $\text{div}(f_0)$ has all integer coefficients (because f_0 is meromorphic), we actually have

$$\text{div}(f_0) + \text{div}(f) \ge 0 \iff \text{div}(f_0) + \lfloor \text{div}(f) \rfloor \ge 0.$$

This means that the dimension satisfies

$$\dim(\mathcal{M}_k(\Gamma)) = \ell(\lfloor \text{div}(f) \rfloor),$$

and we'll now try to understand the right-hand side better. Let ω be the meromorphic differential which pulls back to $f(\tau)(d\tau)^{k/2}$ on $\mathbb{H}$: recall that $\{x_{2,i}\}$ denotes the period 2 elliptic points of $X(\Gamma)$, of which there are a total of ε_2, $\{x_{3,i}\}$ and ε_3 are defined similarly for period 3, and $\{x_i\}$ and ε_∞ denote the cusps.

Definition 227

We define the formal divisor

$$\text{div}(d\tau) = -\sum_i \frac{1}{2} x_{2,i} - \sum_i \frac{2}{3} x_{3,i} - \sum_i x_i.$$

Here, ω kind of looks like $f \cdot (d\tau)^{k/2}$, so we want to "expand out the div."

Proposition 228

We do have

$$\text{div}(\omega) = \text{div}(f) + \frac{k}{2}\text{div}(d\tau).$$

Proof. We just compare coefficients, using the fact that

$$\nu_x(\omega) = \nu_x(f) - \frac{k}{2}\left(1 - \frac{1}{h}\right)$$

at non-cusps, and the same thing without the $\frac{1}{h}$ at cusps. □

If we rearrange this result and take floors term by term (each x-coefficient only appears in one sum, and the first term $\text{div}(\omega)$ has integers),

$$\lfloor \text{div}(f) \rfloor = \text{div}(\omega) + \sum_i \left\lfloor \frac{k}{4} \right\rfloor x_{2,i} + \sum_i \left\lfloor \frac{k}{3} \right\rfloor x_{3,i} + \sum_i \frac{k}{2} x_i.$$

We can now use Riemann-Roch on the floor-div: usually there is a canonical differential term which we want to avoid.

Proposition 229

The degree of $\lfloor \text{div}(f) \rfloor$ is larger than $2g - 2$.

Proof. We know that the divisor is a homomorphism, so we can bash out the sum: we know that $\deg(\text{div}(\omega)) = k(g-1)$, so

$$\deg(\lfloor \text{div}(f) \rfloor) = k(g-1) + \left\lfloor \frac{k}{4} \right\rfloor \varepsilon_2 + \left\lfloor \frac{k}{3} \right\rfloor \varepsilon_3 + \frac{k}{2}\varepsilon_\infty$$

and then we bound by the worst case based on k:

$$\ge \frac{k}{2}(2g-2) + \frac{k-2}{4}\varepsilon_2 + \frac{k-2}{3}\varepsilon_3 + \frac{k}{2}\varepsilon_\infty.$$

This can be rewritten as

$$\boxed{2g - 2 + \frac{k-2}{2}\left(2g - 2 + \frac{\varepsilon_2}{2} + \frac{2\varepsilon_3}{3} + \varepsilon_\infty\right) + \varepsilon_\infty},$$

where the parenthetical term is at least 0 from a previous lecture, and ε_∞ is the equivalence classes of ∞, so it is at least 1. This shows that the degree is large enough. □

And now applying Riemann-Roch in this simple case yields

$$\dim(\mathcal{M}_k(\Gamma)) = \ell(\lfloor \operatorname{div}(f) \rfloor) = \deg(\lfloor \operatorname{div}(f) \rfloor) - g + 1,$$

meaning that

$$\boxed{\dim(\mathcal{M}_k(\Gamma)) = (k-1)(g-1) + \left\lfloor \frac{k}{4} \right\rfloor \varepsilon_2 + \left\lfloor \frac{k}{3} \right\rfloor \varepsilon_3 + \frac{k}{2}\varepsilon_\infty}.$$

We'll now talk briefly about cusp forms, but the story is basically the same: we just have that

$$S_k(\Gamma) \cong L\left(\left\lfloor \operatorname{div}(f) - \sum_i x_i \right\rfloor\right)$$

The strategy is basically the same, except that we require divisors at cusps to have coefficients at least 1. This time, when we calculate the degree of $\lfloor \operatorname{div}(f) \rfloor$, we lose the last ε_∞ term in the boxed expression above, and we need $k \geq 4$ so that the parenthetical term is now a strict "larger than 0." (For the $k = 2$ term, $\lfloor \operatorname{div}(f) \rfloor - \sum_i x_i$ is just the canonical divisor $\operatorname{div}(\lambda)$, so the linear space has dimension g.)

Finally, let's look at the $k \leq 0$ case: $\mathcal{M}_0(X(\Gamma))$ must be constant by Liouville's theorem, and $S_0(X(\Gamma)) = \{0\}$ is the only cusp form. Notice that for $k < 0$, if there is any function $f \in \mathcal{M}_k(\Gamma)$, then $f^{12}\Delta^{-k}$ would be in $S_0(\gamma)$, so indeed we also just have $\mathcal{M}_k(\Gamma) = \{0\}$ for all $k < 0$ (which implies that the spaces of cusps forms are also zero).

So all of our results can be summarized in a compact form:

Theorem 230

For even k, we have

$$\dim(\mathcal{M}_k(\Gamma)) = \begin{cases} (k-1)(g-1) + \left\lfloor \frac{k}{4} \right\rfloor \varepsilon_2 + \left\lfloor \frac{k}{3} \right\rfloor \varepsilon_3 + \frac{k}{2}\varepsilon_\infty & k \geq 2 \\ 1 & k = 0 \\ 0 & k < 0 \end{cases}$$

and

$$\dim(\mathcal{M}_k(\Gamma)) = \begin{cases} (k-1)(g-1) + \left\lfloor \frac{k}{4} \right\rfloor \varepsilon_2 + \left\lfloor \frac{k}{3} \right\rfloor \varepsilon_3 + \left(\frac{k}{2} - 1\right) \varepsilon_\infty & k \geq 4 \\ g & k = 2 \\ 0 & k = 0. \end{cases}$$

This theorem is nice because it gives us results in terms of congruence subgroups which we know how to calculate. And note that when $\Gamma = SL_2(\mathbb{Z})$, we already have a similar result which can be easily derived by the results we know.

Example 231

We can prove that

$$\mathcal{M}_2(\Gamma_0(p)) = S_2(\Gamma_0(p)) \oplus \mathbb{C}G_{2,p}$$

by showing that the dimensions add up.

Basically, S_2 is just missing a dimension of 1 from $\mathcal{M}_2$, and note that $\mathbb{C}G_{2,p} = G_2(\tau) - pG_2(p\tau)$ is an element of $\mathcal{M}_2$ but not an element of $\mathcal{S}_2$.

Example 232

For all k even such that $k(N+1) = 24$, we have $\dim(S_k(\Gamma_0(N))) = 1$.

This in particular means that $S_k(\Gamma_0(N)) = S_k(\Gamma_1(N)) = \mathbb{C}\phi_k$, where

$$\phi_k(\tau) = \eta(\tau)^k \eta(N\tau)^k$$

for the Dedekind eta function η. In particular, this means that

$$\mathcal{M}_2(\Gamma_0(11)) = \mathbb{C}\phi_2 \oplus \mathbb{C}G_{2,p} :$$

we can find some very explicit descriptions of modular forms.

Finally, we have the isomorphism

$$S_k(SL_2(\mathbb{Z})) \cong S_2(\Gamma_0(p)),$$

where $k = p + 1$ and p is an odd prime. And when $p = 11$, we have the simple isomorphism where we multiply by the function $\frac{\phi_2}{\Delta}$.

Diamond and Shurman 3.6 – Vanshika Jain

We'll prove the dimension formulas for the odd-k case here. Here's the result we're trying to show:

Theorem 233

Let Γ be a congruence subgroup of $SL_2(\mathbb{Z})$, and let k be an odd integer. If $-I \in \Gamma$, then $\mathcal{M}_k(\Gamma) = S_k(\Gamma) = 0$. Otherwise, let g be the genus of $X(\Gamma)$, ε_3 be the number of elliptic points of period 3, $\varepsilon_\infty^{\text{reg}}$ be the number of regular cusps and $\varepsilon_\infty^{\text{irr}}$ be the number of irregular cusps. Then

$$\dim M_k(\Gamma) = \begin{cases} (k-1)(g-1) + \left\lfloor \frac{k}{3} \right\rfloor \varepsilon_3 + \frac{k}{2}\varepsilon_\infty^{\text{reg}} + \frac{k-1}{2}\varepsilon_\infty^{\text{irr}} & k \geq 3 \\ 0 & k < 0 \end{cases},$$

$$\dim S_k(\Gamma) = \begin{cases} (k-1)(g-1) + \left\lfloor \frac{k}{3} \right\rfloor \varepsilon_3 + \frac{k-2}{2}\varepsilon_\infty^{\text{reg}} + \frac{k-1}{2}\varepsilon_\infty^{\text{irr}} & k \geq 3 \\ 0 & k < 0 \end{cases}.$$

When $k = 1$, if $\varepsilon_\infty^{\text{reg}} > 2g - 2$, then the dimensions of M_1, S_1 are $\varepsilon_\infty^{\text{reg}}/2$ and 0 respectively, and otherwise, they are at least $\varepsilon_\infty^{\text{reg}}/2$ and equal to $\varepsilon_\infty^{\text{reg}}/2$, respectively.

We'll use the following tools:

- The fact that $\dim(\mathcal{M}_k(\Gamma)) = \ell(\lfloor \text{div}(f) \rfloor)$ still holds for odd k.
- We have an isomorphism of vector spaces $a_n(\Gamma) \to \Omega^{\otimes n/2} X(\Gamma))$. In particular, we can take $f \in A_k(\Gamma)$ and consider f^2 for $n = 2k$.

First, note that $\mathcal{M}_k(\Gamma) = 0$ for all $k < 0$, so $S_k(\Gamma) = 0$ for all $k < 0$ as well. And if k is odd and $-I \in \Gamma$, then

$$f(\tau) = (f[-I]_k)(\tau) = -f(\tau),$$

which implies that the whole space $\mathcal{A}_k(\tau) = 0$, which means that $\mathcal{M}_k(\Gamma) = S_k(\Gamma) = 0$.

From here on out, k will be odd and positive, and we can assume $-I \not\in \Gamma$. Recall that $X(\Gamma)$ can have both regular and irregular cusps: a **irregular cusp** $\pi(s) \in X(\Gamma)$ is one where the group $(\alpha^{-1}\Gamma\alpha)$ is generated by $-\begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix}$, where $\alpha(\infty) = s$. (The cusp is regular if there's no negative sign.) Then we have that the order of vanishing is

$$\nu_{\pi(s)}(f) = \begin{cases} \nu_s(f)/2 & \text{irregular cusp for } s \\ \nu_s(f) & \text{otherwise} \end{cases}$$

for a cusp $\pi(s) \in X(\Gamma)$. (Basically, this is because the Fourier series width is $2h$ instead of just h.) Also, recall that there is an isomorphism between automorphic forms and differentials when $k \in \mathbb{N}$ is even (see above). We'll use that isomorphism now: take ω to be the differential $\omega(f^2)$, which pulls back to $f(\tau)^2(d\tau)^k$ on $\mathbb{H}$. We'll also define $\{x_{3,i}\}, \{x_i\}, \{x_i'\}$ to be the period 3 elliptic points, regular cusps, and irregular cusps, with orders $\varepsilon_3, \varepsilon_\infty^{\text{reg}}$, and $\varepsilon_\infty^{\text{irr}}$.

We'll similarly define the formal divisor

$$\text{div}(d\tau) = -\sum_i \frac{2}{3} x_{3,i} - \sum_i x_i - \sum_i x_i'.$$

Note that because $-I \not\in \Gamma$, there are no period 2 elliptic points, so this is consistent with last time's definition. Replace f and $\frac{k}{2}$ from the last lecture to show similarly that

$$\text{div}(\omega) = 2\text{div}(f) + k\text{div}(d\tau),$$

and now we can combine these two to find that

$$\text{div}(f) = \frac{1}{2}\text{div}(\omega) + \sum_i \frac{k}{3} x_{3,i} + \sum_i \frac{k}{2} x_i + \sum_i \frac{k}{2} x_i'.$$

Taking pointwise floor, the order of vanishing is integral everywhere except at $x_{3,i}, x_i, x_i'$. For the elliptic points x of order 3, we know that we can write the order of vanishing as

$$\nu_x(f) = m + \frac{j}{3} \implies \frac{1}{2} v_x(\omega) = m + \frac{j-k}{3},$$

but $j \equiv k \bmod 3$ because $\frac{1}{2}v_x(\omega)$ is integral, so the integral part of the order of vanishing is just $\frac{1}{2}v_x(\omega) + \lfloor \frac{k}{3} \rfloor$. Similarly, we can find the integral parts at x_i and x_i': these have additional terms of $\frac{k}{2}$ and $\frac{k-1}{2}$ (plus the original $\frac{1}{2}v_x(\omega)$), so

$$\dim(M_k(\Gamma)) = \ell(\lfloor \dim(f) \rfloor) = \ell\left(\frac{1}{2}\text{div}(\omega) + \sum_i \left\lfloor \frac{k}{3} \right\rfloor x_{3,i} + \sum_i \frac{k}{2} x_i + \sum_i \frac{k-1}{2} x_i' \right).$$

Our goal is to use a corollary of Riemann-Roch, so we want to get a bound on the degree of the divisor here (specifically, we again want it to be at least $2g-2$). But

$$\deg(\lfloor \text{div}(f) \rfloor) = k(g-1) + \left\lfloor \frac{k}{3} \right\rfloor \varepsilon_3 + \frac{k}{2}\varepsilon_\infty^{\text{reg}} + \frac{k-1}{2}\varepsilon_\infty^{\text{irr}} > (k-2)\left(g - 1 + \frac{\varepsilon_3}{3} + \frac{\varepsilon_\infty}{2}\right) + (2g-2) > 2g-2$$

for all $k \geq 3$, so now the simpler form of Riemann-Roch tells us that

$$\ell(\lfloor \text{div}(f) \rfloor) = (k-1)(g-1) + \left\lfloor \frac{k}{3} \right\rfloor \varepsilon_3 + \frac{k}{2}\varepsilon_\infty^{\text{reg}} + \frac{k-1}{2}\varepsilon_\infty^{\text{irr}},$$

as desired. For cusp forms, we'll look at functions of the form $f_0 f \in \mathcal{A}_k(\Gamma)$: note that at a regular cusp,

$$v_x(f_0 f) > 0 \iff v_x(f_0 f) \geq 1 \iff v_x(f_0) + v_x(f) - 1 \geq 0,$$

and at an irregular cusp,

$$v_x(f_0 f) > 0 \iff v_x(f_0 f) \geq v_x(f_0) + v(f) - \frac{1}{2} \geq 0.$$

(this is coming from the order of vanishing from earlier in the lecture). So we want to measure the dimension

$$\ell\left(\lfloor \operatorname{div}(f) - \sum_i x_i - \sum \frac{1}{2} x_i' \rfloor\right),$$

and we just do similar calculations to get the result in our main theorem.

To summarize, we used the isomorphism with f^2 instead of f, which gives us our differential. This allows us to calculate an explicit formula for the divisor $\operatorname{div}(f)$, and then we do some case-by-case analysis to get the integral parts.

We'll now prove that $A_k(\Gamma)$ does contain a nonzero f when $k > 0$ is odd and $-I \notin \Gamma$. Take a nonzero differential $\lambda \in \Omega^1(X(\Gamma))$ – for example, we can consider the divisor that pulls back to $j'(\tau)d\tau$. Now take an element $x_0 \in X(\Gamma)$ – by Riemann-Roch, the degree of the divisor of λ is $2g - 2$, so

$$\operatorname{div}(\lambda) - 2(g-1)x_0 \in \operatorname{Div}^0,$$

where Div^0 is the set of divisors of degree 0. (In general, Div^ℓ denotes the divisors of nonzero meromorphic functions on $X(\Gamma)$.) Abel's theorem tells us that we have a mapping $\operatorname{Div}^0/\operatorname{Div}^\ell \to \mathbb{C}^g/\Lambda g$: let $x + \Lambda g$ be the image of the element $\operatorname{div}(\lambda) - 2(g-1)x_0 + \operatorname{Div}^\ell$. If a divisor $D \in \operatorname{Div}^0$ has that $D + \operatorname{Div}^\ell$ maps to $\frac{z}{2} + \Lambda g$ (where $z \in \mathbb{C}^g$), then we know that

$$2D = \operatorname{div}(\lambda) + 2(g-1)x_0 + \operatorname{div}(f)$$

for some $f \in \mathbb{C}(X(\Gamma)$. We can rearrange this to

$$\operatorname{div}(f\lambda) = 2(D + (g-1)x_0),$$

and remember our goal now is to find a automorphic form f which is weight-k and nonzero, and we'll do this by finding something of weight 1 and raise it to the kth power. **Let $\tilde{f}$ be the function such that $f\lambda$ on $X(\Gamma)$ pulls back to $\tilde{f}(\tau)d\tau$ on $\mathbb{H}$.** This is nice because we know what $d\tau$ is: we find that

$$\operatorname{div}(\tilde{f}) = \operatorname{div}(f\lambda) - \operatorname{div}(d\tau) = 2(D + (g-1)x_0) + \sum_i \frac{2}{3} x_{3,i} + \sum_i x_i + \sum_i x_i'.$$

We know that at elliptic points of period 3, $v_\tau(\tilde{f})$ is $3v_{\pi(\tau)}(\tilde{f})$ and we have a 2 in the numerator for this term, so this is even. Since $\tilde{f}$ is weight-2 invariant, there exists a function f_1 such that $f_1^2 = \tilde{f}$ and $f_1[\gamma]_1 = \chi(\gamma)f_1$, where $\chi : \Gamma \to \{\pm 1\}$ (so it's an automorphic form up to some negative sign). Let Γ' be the set such that $\chi(\gamma) = 1$. If the index is 1, then we're good – otherwise, we find a function f_2 which is negative at the correct spots, and multiplying f_1 by f_2 means we have our element of $A_k(\Gamma)$, as desired.

19 April 21, 2020

Diamond and Shurman 3.8, 4.1 – Michelle Xu

We'll cover **cusps** for congruence subgroups (explicit forms, cusps in relation to double cosets) and **Eisenstein series** (review of material from Serre) in this lecture: the topics might seem disjoint, but we'll connect them in the next lecture.

Recall that a **cusp** is an equivalence class of $\mathbb{Q} \cup \{\infty\}$ under action by Γ: this is also denoted $\Gamma \backslash (\mathbb{Q} \cup \{\infty\})$. We know that there is always an element of $SL_2(\mathbb{Z})$ which maps ∞ to r for any rational r, so there is only one cusp for $SL_2(\mathbb{Z})$. However, this won't be true with our congruence subgroups.

Lemma 234

Let $s = \frac{a}{c}$ and $s' = \frac{a'}{c'}$ be elements of $\mathbb{Q} \cup \{\infty\}$ such that $\gcd(a, c) = \gcd(a', c') = 1$ are coprime. Then

$$s' = \gamma(s) \iff \begin{bmatrix} a' \\ c' \end{bmatrix} = \pm\gamma \begin{bmatrix} a \\ c \end{bmatrix}.$$

Proof. When c, c' are both nonzero, we know that if $\gamma = \begin{bmatrix} p & q \\ r & t \end{bmatrix}$,

$$s' = \gamma(s) \iff \frac{a'}{c'} = \frac{pa + qc}{ra + tc} = \frac{a}{c}.$$

But both fractions are in lowest terms because we can invert γ^{-1}, so we must have $a' = pa + qc, c' = ra + tc$ as desired.

If $c = 0$, then $a = 1$ based on our condition, so $\frac{a'}{c'} = \frac{p}{r}$. Here, we have $\gcd(p, r) = 1$ because $pt - qr = 1$, which means that $a' = p, c' = r$. (And $c' = 0$ is similar.) □

We'll now focus more on the congruence subgroups:

Lemma 235

If $\gamma \in \Gamma(N)$, then

$$\begin{bmatrix} a' \\ c' \end{bmatrix} = \gamma \begin{bmatrix} a \\ c \end{bmatrix} \iff \begin{bmatrix} a' \\ c' \end{bmatrix} \equiv \gamma \begin{bmatrix} a \\ c \end{bmatrix} \mod N.$$

Proof. Here, γ is the identity matrix mod N, so left-to-right is clear. For the other direction, we start with the specific case $a = 1, c = 0$, which means that $a' = 1 \mod N$. Bezout tells us that for any $x, y \in \mathbb{Z}$, we have $m, n \in \mathbb{Z}$ such that $mx + ny = \gcd(x, y)$, and in general we can make $mx + ny$ equal to any multiple of $\gcd(x, y)$. In our specific case, because a' and c' are coprime, we can find β, δ such that

$$a'\beta + c'\delta = \frac{1 - a}{N},$$

and this means we can use the matrix

$$\gamma = \begin{bmatrix} a' & \beta N \\ c' & 1 + \delta N \end{bmatrix} \in \Gamma(N).$$

In general (when we don't have $(a, c) = (1, 0)$), we know that there exist $b, d \in \mathbb{Z}$ such that $ad - bc = 1$, so using the matrix $\alpha = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

$$\alpha \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} a \\ c \end{bmatrix} \implies \alpha^{-1} \begin{bmatrix} a \\ c \end{bmatrix} \equiv \begin{bmatrix} 1 \\ 0 \end{bmatrix} \mod N,$$

so we can first find γ' such that

$$\alpha^{-1} \begin{bmatrix} a' \\ c' \end{bmatrix} = \gamma' \alpha^{-1} \begin{bmatrix} a \\ c \end{bmatrix},$$

and use $\gamma = \alpha\gamma'\alpha^{-1}$ because $\Gamma(N)$ is a **normal subgroup** of $SL_2(\mathbb{Z})$. □

We'll use these lemmas to prove more about our cusps:

Proposition 236

Take the same notation as above, so $s = \frac{a}{c}, s' = \frac{a'}{c'}$. Then we have an explicit description for s and s' to correspond to the same cusp:

$$\Gamma(N)s' = \Gamma(N)s \iff \begin{bmatrix} a' \\ c' \end{bmatrix} = \pm \begin{bmatrix} a \\ c \end{bmatrix} \bmod N,$$

$$\Gamma_1(N)s' = \Gamma_1(N)s \iff \begin{bmatrix} a' \\ c' \end{bmatrix} = \pm \begin{bmatrix} a + jc \\ c \end{bmatrix} \bmod N,$$

$$\Gamma_0(N)s' = \Gamma_0(N)s \iff \begin{bmatrix} ya' \\ c' \end{bmatrix} = \pm \begin{bmatrix} a + jc \\ yc \end{bmatrix} \bmod N$$

for some j, y.

Proof. For $\Gamma(N)$, the left hand side is equivalent to saying that $s' = \gamma(s)$ for some $\gamma \in \Gamma(N)$, which is equivalent to $\begin{bmatrix} a' \\ c' \end{bmatrix} = \pm\gamma \begin{bmatrix} a \\ c \end{bmatrix}$ by the first lemma, and then we can apply the second lemma.

For $\Gamma_1(N)$, any matrix can be written in the form

$$\begin{bmatrix} 1 & * \\ 0 & 1 \end{bmatrix} \bmod N = \begin{bmatrix} aN + 1 & bN + j \\ cN & dN + 1 \end{bmatrix}.$$

We decompose the congruence subgroup as

$$\Gamma_1(N) = \bigcup_j \Gamma(N) \begin{bmatrix} 1 & j \\ 0 & 1 \end{bmatrix},$$

and we can now apply the results we already know for Γ: the left hand side is equivalent to saying that $s' \in \Gamma_1(N)s$, so $s' \in \Gamma(N) \begin{bmatrix} 1 & j \\ 0 & 1 \end{bmatrix} s$ for some j. But then we can just multiply out and apply the lemmas for $\Gamma(N)$, and this tells us gives us the desired result.

For Γ_0, we'll decompose again:

$$\Gamma_0(N) = \bigcup_y \Gamma_1(N) \begin{bmatrix} x & k \\ N & y \end{bmatrix},$$

where we take the union over y relatively prime to N, and we need $xy - Nk = 1$. And the rest is a similar calculation: the left side is equivalent to $s' \in \Gamma_1(N) \begin{bmatrix} x & k \\ N & y \end{bmatrix} x$ for some y, which is equivalent to $\Gamma_1(N)s' = \Gamma_1(N)\frac{xa+kc}{Na+yc}$, and applying the result for Γ_1 tells us that

$$\begin{bmatrix} a' \\ c' \end{bmatrix} = \pm \begin{bmatrix} xa + kc + jyc \\ yc \end{bmatrix} \bmod N,$$

and multiply a factor of y on both sides to get the result. □

This means we have explicit conditions for checking whether two elements in $\mathbb{Q} \cup \{\infty\}$ are the same cusps, and this allows us to count the number of cusps for each subgroup.

Instead, we'll discuss these cusps in a group-theoretic manner:

Definition 237

Let G be an arbitrary group with subgroups H_1, H_2. A **double coset** of G is a subset of the form $H_1 g H_2$, and the space of double cosets is denoted $H_1 \backslash G / H_2$.

These double cosets are disjoint, and we can therefore decompose as

$$G = \bigcup_{g \in R} H_1 g H_2.$$

Proposition 238

Let Γ be a congruence subgroup of $SL_2(\mathbb{Z})$. Let P be the **parabolic subgroup**

$$P = \left\{ \pm \begin{bmatrix} 1 & j \\ 0 & 1 \end{bmatrix} : j \in \mathbb{Z} \right\}.$$

Then the map between $\Gamma \backslash SL_2(\mathbb{Z})/P$ and cusps of Γ is a bijection: $\Gamma \begin{bmatrix} a & b \\ c & d \end{bmatrix} P$ maps to $\Gamma(a/c)$.

(We won't prove this, but it's a nice way to study cusps and is more general.)

We'll finish by reviewing the **Eisenstein series**: recall that we define these to be

$$G_k(\tau) = \sum_{(c,d) \in \mathbb{Z}}' \frac{1}{(c\tau + d)^k},$$

where we sum over all pairs of integers except $(0, 0)$. We also define the **normalized Eisenstein series** $E_k(|tau) = \frac{G_k(\tau)}{2\zeta(k)}$.

We'll be rewriting the Eisenstein series in nicer forms now:

Lemma 239

We have

$$G_k(\tau) = \zeta(k) \sum_{\gcd(c,d)=1} \frac{1}{(c\tau + d)^k}.$$

Proof. Reorganize the order of summation by the gcd (absolute convergence allows us to do this for $k \geq 4$):

$$G_k(\tau) = \sum_{n=1}^{\infty} \sum_{\gcd(c,d)=n} \frac{1}{(c\tau + d)^k} = \sum_{n=1}^{\infty} \frac{1}{n^k} \sum_{\gcd(c,d)=1} \frac{1}{(c\tau + d)^k},$$

which is what we want. □

Lemma 240

Let P_+ be the positive part of the parabolic subgroup of $SL_2(\mathbb{Z})$ (P, except without the $\pm$). Then

$$E_k(\tau) = \frac{1}{2} \sum_{\gamma \in P_+ \backslash SL_2(\mathbb{Z})} j(\gamma, \tau)^{-k},$$

where $j(\gamma, \tau) = (c\tau + d)$.

Proof. Take the previous lemma and divide to find

$$E_k(\tau) = \sum_{\gcd(c,d)=1} \frac{1}{(c\tau+d)^k}.$$

We know that c, d have no common divisor, so we can index with $SL_2(\mathbb{Z})$ as long as we remove duplicates of the same (c, d). And this requires proving that the sets of $SL_2(\mathbb{Z})$ with the same (c, d) are exactly the same as the orbits of P_+. One direction is easy – we can check that multiplying $SL_2(\mathbb{Z})$ by P_+ preserves the (c, d) pair. For the other direction, if we consider both $\begin{bmatrix} a_1 & b_1 \\ c & d \end{bmatrix}$ and $\begin{bmatrix} a_2 & b_2 \\ c & d \end{bmatrix}$, then

$$a_1 = a_2 + \frac{(b_1 - b_2)c}{d},$$

and indeed $n = \frac{b_1-b_2}{d}$ means we have a perfect indexing of (c, d) pairs. Substituting everything in gives the result. □

This in particular show that $E_k(\tau)$ is weakly modular of weight k: this is because we act with some γ and use properties of the factor of automorphy $j(\gamma, \tau)$, as well as the fact that multiplying by γ' onto some $\gamma \in P_+ \backslash SL_2(\mathbb{Z})$ gives another element of that equivalence class.

We found in Serre that

$$\mathcal{M}_k = \mathcal{S}_k \oplus \mathbb{C}G_k,$$

which motivates the following definition in general:

Definition 241

The **weight-*k* Einstein space** $\mathcal{E}_k(\Gamma)$ of a congruence subgroup Γ is the quotient space $\mathcal{M}_k/\mathcal{S}_k$.

We will prove later that this is a subspace of $\mathcal{M}_k$ and also complementary to $\mathcal{S}_k$.

Proposition 242

The dimension of the Eisenstein space $\mathcal{E}_k(\Gamma)$ is ε_∞ when $k \geq 4$ is even, $\varepsilon_\infty^{\text{reg}}$ when $k \geq 3$ is odd and $-I \notin \Gamma$, $\varepsilon_\infty - 1$ when $k = 2$, $\varepsilon_\infty^{\text{reg}}$ when $k = 1$ and $-I \notin \Gamma$, 1 when $k = 0$, and 0 otherwise.

This is basically just casework from previous lectures and subtracting dimensions.

Diamond and Shurman 4.2 – Andrew Gu

In this lecture, we will generalize the construction from last section to construct multiple Eisenstein series for congruence subgroups Γ, and we'll get an explicit basis for the Eisenstein space. Throughout this lecture, overline means reduction mod N, $P_+^N = P_+ \cap \Gamma(N)$, $k \geq 3$ is a positive integer, and $\overline{v} \in (\mathbb{Z}/n\mathbb{Z})^2$ is a **row vector of order** N (in other words, $\gcd(c, d, N)$ is 1 for any lift $v = (c, d)$ of $\overline{v}$). Also, let ϵ_N be equal to $\frac{1}{2}$ for $N = 1$ or 2 and be equal to 1 otherwise (this is compensating for the $-I$ being in $\Gamma(N)$).

Definition 243

For any row vector $\overline{v}$, define the Eisenstein series

$$E_k^{\overline{v}}(\tau) = \epsilon_N \sum_{\substack{(c,d)\equiv v \bmod N \\ \gcd(c,d)=1}} (c\tau+d)^{-k}.$$

Proposition 244

Let δ be any element of $SL_2(\mathbb{Z})$ where the bottom row of the matrix reduces to $\overline{(c_v, d_v)}$ mod N. Then we can group matrices by bottom row:

$$E_k^{\overline{v}}(\tau) = \epsilon_N \sum_{\gamma \in P_+^N \backslash \Gamma(N)\delta} j(\gamma, \tau)^{-k},$$

where $j(\gamma, \tau)$ is the usual factor of automorphy $c\tau + d$.

We'll omit this proof for now. Our goal is to show that this series is an element of $\mathcal{M}_k(\Gamma(N))$: to do that, we need to show that it is holomorphic, weight-k invariant, and holomorphic at all cusps. The first property is basically identical to the $N = 1$ case (because of uniform convergence, the function is again holomorphic everywhere).

To understand whether the transformation law works correctly, we consider the weight-k operator:

Proposition 245 (Transformation law)

For any $\gamma \in SL_2(\mathbb{Z})$, we have

$$(E_k^{\overline{v}}[\gamma]_k)(\tau) = E_k^{\overline{v\gamma}}(\tau).$$

In other words, we just end up with another Eisenstein series corresponding to a different row vector.

Proof. This is just a computation: recall that the weight-k operator is defined via

$$(f[\gamma]_k)(\tau) = j(\gamma, \tau)^{-k} f(\gamma(\tau)),$$

so the left side is equal to

$$\epsilon_N j(\gamma, \tau)^{-k} \sum_{\gamma' \in P_+^N \backslash \Gamma(N)\delta} j(\gamma', \gamma(\tau))^{-k}$$

which can be simplified via properties of j to

$$= \epsilon_N \sum_{\gamma' \in P_+^N \backslash \Gamma(N)\delta} j(\gamma'\gamma, \tau)^{-k},$$

and relabeling indices, this is summing $j(\gamma'', \tau)^{-k}$ over $\gamma' \in P_+^N \backslash \Gamma(N)\delta\gamma$, which is exactly the right side. □

Corollary 246

$E_k^{\overline{v}}$ is weight-k invariant.

Proof. For any $\gamma \in \Gamma(N)$, $\overline{v\gamma} = \overline{v}$ (because γ is the identity mod N), so this follows directly from the above result:

$$(E_k^{\overline{v}}[\gamma]_k)(\tau) = E_k^{\overline{v\gamma}}(\tau),$$

as desired. □

It remains to show that the behavior at the cusps is what we want: we'll start by looking at the cusp ∞. The only terms in the Eisenstein series that do not vanish are those where $(c\tau + d)^{-k}$ does not go to zero, so we must have $c = 0$ and $|d| = 1$, and this only occurs when $\overline{v} = \pm\overline{(0, 1)}$. $N = 1$ and 2 are special cases, because we have both $(0, 1)$ and $(0, -1)$. When k is odd, these contributions cancel, and otherwise we get an extra contribution of $\frac{1}{2}((-1)^k + (-1)^k) = (-1)^k$ from the two $j(\gamma, \tau)$ terms. Thus, as $\text{Im}(\tau) \to \infty$, $E_k^{\overline{v}}(\tau)$ is only nonzero when $\overline{v} = \pm\overline{(0, 1)}$, and furthermore we must have either $N \geq 3$ or k even.

Remark 247. *This is part of why the normalization ϵ_N factor is nice: the values of our normalized Eisenstein series $E_k^{\overline{v}}$ are always integers at cusps.*

Now note that when k is odd and $n \in \{1, 2\}$, the Eisenstein space is zero (because $-I$ is in the congruence subgroup). So we'll assume that we're working with $N \geq 3$ and k even until we do our Fourier coefficient calculation.

Proposition 248

$E_k^{\overline{v}}$ is holomorphic at all cusps.

Proof. Let $\overline{v} = \overline{(c, d)}$ be the bottom row of $\delta \in SL_2(\mathbb{Z})$. If $s = \frac{a'}{c'}$ is any cusp in $\mathbb{Q} \cup \{\infty\}$, let $\alpha = \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix} \in SL_2(\mathbb{Z})$ take ∞ to s: then the transformation law in Proposition 245 tells us that

$$E_k^{\overline{v}}(s) = E_k^{\overline{v}}[\alpha]_k(\infty) = E_k^{\overline{v\alpha}}(\infty),$$

where the superscript can be written as $\overline{(0, 1)\delta\alpha}$. We know that this last Eisenstein series is only nonzero if

$$\overline{(0, 1)\delta\alpha} = \pm\overline{0, 1} \implies \begin{bmatrix} a' \\ c' \end{bmatrix} \equiv \pm \begin{bmatrix} -d \\ c \end{bmatrix} \text{ mod } N.$$

This means that $E_k^{\overline{v}}$ is only nonvanishing at the cusp (orbit) $\Gamma(N)(-d/c)$ and is zero everywhere else. Thus the value is always finite, so we do have holomorphicity at the cusps. □

Note now that all cusps are regular, so $\mathcal{E}_k(\Gamma(N))$ has dimension $\varepsilon_\infty(\Gamma(N))$ (since we're assuming $k \geq 3$ and N is not 1 or 2). Also notice above that each cusp can be associated with an explicit Eisenstein series which is only nonzero at that cusp, so we can use that set of Eisenstein series to create a basis.

Corollary 249

We can construct an Eisenstein series by summing over cosets: for **any congruence subgroup** Γ of level N, the sum over coset representatives

$$E_{k,\Gamma}^{\overline{v}} = \sum_{\gamma_J \in \Gamma(N) \backslash \Gamma} E_k^{\overline{v}}[\gamma_j]_k$$

is an element of $\mathcal{M}_k(\Gamma)$.

We'll now move on to computing Fourier series. $E_k^{\overline{v}}$ is normalized, but we want to instead look at the non-normalized series (where we don't require $\gcd(c, d) = 1$)

$$G_k^{\overline{v}}(\tau) = \sum_{(c,d) \equiv v \text{ mod } n}' (c\tau + d)^{-k}$$

when calculating these Fourier coefficients. We'll group our terms according to $\gcd(c, d)$, which is always coprime to N by assumption, and this yields

$$G_k^{\overline{v}}(\tau) = \sum_{\substack{n=1 \\ \gcd(n,N)=1}} \sum_{\substack{(c,d) \equiv v \text{ mod } N \\ \gcd(c,d)=N}} (c\tau + d)^{-k},$$

and now factoring out the gcd term in the inner sum yields

$$= \sum_{\substack{n=1 \\ \gcd(n,N)=1}} \frac{1}{n^k} \sum_{\substack{(c',d') \equiv n^{-1}v \text{ mod } N \\ \gcd(c',d')=1}} (c'\tau + d')^{-k}.$$

Thus, we can write

$$G_k^{\overline{v}}(\tau) = \frac{1}{\epsilon_N} \sum_{n \in (\mathbb{Z}/n\mathbb{Z})^*} \zeta_+^n(k) E_k^{n^{-1}\overline{v}}(\tau),$$

where

$$\zeta_+^n(k) = \sum_{\substack{m=1 \\ m \equiv n \bmod N}}^{\infty} \frac{1}{m^k}$$

is the **modified zeta function** for any $n \in (\mathbb{Z}/n\mathbb{Z})^*$. (This reduces to the usual zeta function for $N = 1$.) We can similarly rewrite the E series in terms of the G series:

Proposition 250

We have

$$E_k^{\overline{v}}(\tau) = \sum_{n \in (\mathbb{Z}/n\mathbb{Z})^*} \zeta_+^n(k, \mu) G_k^{n^{-1}\overline{v}}(\tau)$$

for the function

$$\zeta_+^n(k, \mu) = \sum_{\substack{m=1 \\ m \equiv n \bmod N}}^{\infty} \frac{\mu(m)}{m^k}$$

where $\mu(m)$ is 0 for non-squarefree m, and otherwise $(-1)^k$ if k prime numbers divide m.

We prove this with a lot of calculation and a Möbius inversion: one main idea is that $\sum_{d|m} \mu(d) = 1$ when $m = 1$ and 0 otherwise. And now that recalling that we can choose a collection of vectors that represent each cusp once, giving a basis of $\{E_k^{\overline{v}}\}$s: the result above now tells us that we can get a basis of $\{G_k^{\overline{v}}\}$s as well.

To get a Fourier series, we'll need the following lemma which was proved in Serre earlier in this class:

Lemma 251

For any $k \geq 2$ and $\tau \in \mathbb{H}$, we have

$$\sum_{d \in \mathbb{Z}} \frac{1}{(\tau + d)^k} = C_k \sum_{m=1}^{\infty} m^{k-1} q^m,$$

where $q = e^{2\pi i \tau}$ and $C_k = \frac{(-2\pi i)^k}{(k-1)!}$.

This now allows us to compute the Fourier series of $G_k^{\overline{v}}$ (remembering that we're summing over all nonzero (c, d)) by rewriting as

$$G_k^{\overline{v}}(\tau) = \sum_{c \equiv c_v} \sum_{d \in \mathbb{Z}} \frac{1}{(c\tau + d_v + Nd)^k} = \frac{1}{N^k} \sum_{c \equiv c_v} \sum_{d \in \mathbb{Z}} \frac{1}{\left(\frac{c\tau + d_v}{N} + d\right)^k}.$$

Notice that this does sum over all $(c, d) \equiv (c_v, d_v) \bmod N$, and now we'll break this up by values of c.

- $c = 0$ yields the constant term of the Fourier series (the above lemma will show that all other values of c have no constant term, or alternatively we can think of taking τ to $+i\infty$): if $\overline{c_v} = 0$, then this is

  $$\sum_{d \equiv d_v}{}' \frac{1}{d^k},$$

  and otherwise it is zero.
- For $c > 0$ (corresponding to points in the upper half plane), we can use our lemma with $\frac{c\tau + d_v}{N}$ in place of τ.

This then evaluates to

$$\frac{1}{N^k}\sum_{\substack{c\equiv c_v\\c>0}}\sum_{d\in\mathbb{Z}}\frac{1}{\left(\frac{c\tau+d_v}{N}+d\right)^k}=\frac{C_k}{N^k}\sum_{\substack{c\equiv c_v\\c>0}}\sum_{m=1}^{\infty}m^{k-1}\mu_N^{d_v m}q_N^{cm},$$

where $\mu_N = e^{2\pi i/N}$ and $q = e^{2\pi i\tau/N}$. (The c and d_v in the exponents can be understood if we expand out $e^{2\pi i(c\tau+d_v)/N}$ as in the lemma.) This can then be rewritten by summing over constant exponents of q_N as

$$=\frac{C_k}{N^k}\sum_{\substack{c\equiv c_v\\c>0}}\sum_{\substack{m|n\\n/m\equiv c_v\\m>0}}^{\infty}m^{k-1}\mu_N^{d_v m}q_N^{n},$$

and we have a Fourier expansion for the $c > 0$ terms.

- On the other hand, for $c < 0$, we can write $\frac{c\tau+d_v}{N}+d=-\left(-\frac{c\tau-d_v}{N}-d\right)$. The first term in the parentheses, $-\frac{c\tau-d_v}{N}$, now does exist in the upper half-plane, and applying the lemma gives something similar: we have

  $$\frac{1}{N^k}\sum_{\substack{c\equiv c_v\\c<0}}\sum_{d\in\mathbb{Z}}\frac{1}{\left(\frac{c\tau+d_v}{N}+d\right)^k}=(-1)^k\frac{C_k}{N^k}\sum_{\substack{c\equiv c_v\\c>0}}\sum_{m=1}^{\infty}m^{k-1}\mu_N^{-d_v m}q_N^{-cm},$$

  and we can now rewrite this sum again by summing over the exponent of q_N: if we flip signs to force c to now be positive, m must be negative, and the $(-1)^k$ term simplifies with the $(-1)^{k-1}$ from the m exponent. We end up with

  $$\frac{C_k}{N^k}\sum_{\substack{c\equiv c_v\\c>0}}\sum_{\substack{m|n\\n/m\equiv c_v\\m<0}}-m^{k-1}\mu_N^{d_v m}q_N^{n},$$

  where notice that we now sum over the negative divisors of n instead of the positive ones.

Putting all of these together, we get the full Fourier expansion for $G_k^{\overline{v}}$: the q_N^n-coefficient of $G_k^{\overline{v}}(\tau)$ for a point $\overline{v}$ of order n is

$$\frac{C_k}{N^k}\sigma_{k-1}^{\overline{v}}(n),$$

where

$$\sigma_{k-1}^{\overline{v}}(n)=\sum_{\substack{m|n\\n/m\equiv c_v}}\operatorname{sgn}(m)m^{k-1}\mu_N^{d_v m}.$$

As a final note, this Fourier expansion tells us that the coefficients are polynomially-bounded for any $E_k^{\overline{v}}$, and thus we find directly that we have a modular form.

20 April 21, 2020

Diamond and Shurman 4.3 – Andrew Lin

Recall that last week, we introduced the Eisenstein series associated to specific congruence subgroups Γ. We were able to describe Eisenstein series that live in $\mathcal{M}_k(\Gamma)$, and we also found that we could describe the cusp "Eisenstein space"

$$\mathcal{E}_k(\Gamma)=\mathcal{M}_k(\Gamma)/\mathcal{S}_k(\Gamma),$$

computing explicit basis elements of $\mathcal{E}_k(\Gamma(N))$. The point of this section is to start giving some background for looking at modular forms for $\Gamma_1(N)$, as well as cover some general number theory concepts.

Throughout this, we will let G_N denote the multiplicative group $(\mathbb{Z}/N\mathbb{Z})^*$: recall that this group has order $\phi(N)$, and it consists of the residues mod N that are relatively prime to N.

Definition 252

A **Dirichlet character** mod N is a homomorphism $\chi : G_N \to \mathbb{C}^*$.

Here are a few useful properties we can extract from this definition:

- Because G_N is a finite group, all elements have finite order, so $\chi(n)$ is a root of unity for all $n \in G_N$.
- The set of Dirichlet characters form a group, called the **dual group** $\hat{G}_N$ of G_N. Indeed, the law of composition for two characters χ, ψ is the product character $(\chi\psi)(n) = \chi(n)\psi(n)$. This group has an identity element $\mathbf{1}_N$, which just sends everything to 1 (known as the trivial character), and the inverse of a character χ is $\chi^{-1} = \overline{\chi}$, which takes the complex conjugate of $\chi(n)$ for all n.
- Recall that for prime numbers p, the multiplicative group $(\mathbb{Z}/p\mathbb{Z})^*$ is cyclic. Then the dual group is also cyclic: each character is determined by which root of unity a primitive root g is sent to.

This last result can be generalized to all integers, not just primes:

Fact 253

The dual group $\hat{G}_N$ is isomorphic to G_N (and in particular, there are $\phi(N)$ Dirichlet characters mod N).

The proof can be found in Serre: the main idea is to induct on the order of the group by considering subgroups of G_N and noting that for any subgroup H of G, we can decompose G as the product of H and G/H (and same for the dual groups). However, there is no canonical isomorphism in general (for example, we need to pick a primitive root arbitrarily for the prime n case).

The next result generalizes the duality between the two groups:

Proposition 254 (Orthogonality relations)

We have

$$\sum_{n \in G_N} \chi(n) = \begin{cases} \phi(N) & \chi = \mathbf{1}_N \\ 0 & \text{otherwise} \end{cases}$$

and

$$\sum_{\chi \in \hat{G}_N} \chi(n) = \begin{cases} \phi(N) & n = 1 \\ 0 & \text{otherwise.} \end{cases}$$

These may look familiar (for example) from 18.702.

Proof. We'll just prove the first result – the two proofs are basically the same. If $\chi = \mathbf{1}_N$, then we are just adding up the total number of elements in the group, which is clearly $\phi(n)$ by the above fact. Otherwise, there is some element y such that $\chi(y) \neq 1$. Then summing over the whole group,

$$\sum_{n \in G_N} \chi(n) = \sum_{n \in G_N} \chi(ny) = \sum_{n \in G_N} \chi(n)\chi(y),$$

which means that

$$(\chi(y) - 1)(\sum_{n \in G_N} \chi(n)) = 0,$$

and the result follows. We do the same thing for the second relation, except we multiply by a character instead (and use the fact that we still have a group). □

We'll now say a bit more about this idea of building up characters from smaller ones: suppose that d divides N, and we have a Dirichlet character χ_d mod d. Then we can construct a Dirichlet character mod N in the natural way by letting

$$\chi_N(n \text{ mod } N) = \chi_d(n \text{ mod } d),$$

but this is only always consistent if d divides N. We denote this via $\chi_N = \chi \circ \pi_{N,d}$, where $\pi_{N,d}$ takes an element of G_N and reduces it mod d.

Definition 255

The **conductor** of a Dirichlet character χ mod N is the smallest $d|N$ such that we can write $\chi = \chi_d \circ \pi_{N,d}$ for some χ_d.

Because of the simple group structure here, we can equivalently say that the conductor is the smallest d such that χ is identically equal to 1 for all elements in the normal subgroup

$$K = \{n \in G_N : n \equiv 1 \text{ mod } d\},$$

which is the kernel of the projection map.

Example 256

The conductor of the Dirichlet character ψ mod 12 sending $\overline{1}, \overline{7}$ to $+1$ and $\overline{5}, \overline{11}$ to -1 is 3.

Definition 257

A Dirichlet character mod N is **primitive** if its conductor is N.

For example, ψ "only cares about the element mod 3," so it is not primitive.

We can extend the definition of χ to all integers by turning it into a function $\chi : \mathbb{Z} \to \mathbb{C}$:

$$\chi(n) = \begin{cases} \chi(n \text{ mod } N) & \gcd(n, N) = 1 \\ 0 & \text{otherwise.} \end{cases}$$

This function is indeed consistent with the original χ we described, and it is still multiplicative, although it is not a homomorphism because 0 is not invertible in $\mathbb{C}$. (Keep in mind that this also means the trivial character $\mathbf{1}_N$ is no longer just equal to 1 everywhere, and for example $\chi(0)$ is only nonzero when $N = 1$.) Then we can restate the orthogonality relations in a more helpful form for later use, though they are not really any different:

$$\sum_{n=0}^{N-1} \chi(n) = \begin{cases} \phi(N) & \chi = \mathbf{1}_N \\ 0 & \text{otherwise} \end{cases}, \quad \sum_{\chi \in \hat{G}_N} \chi(n) = \begin{cases} \phi(N) & n \equiv 1 \text{ mod } N \\ 0 & \text{otherwise.} \end{cases}$$

Definition 258

The **Gauss sum** of a character χ mod N is

$$g(\chi) = \sum_{n=0}^{N-1} \chi(n)\mu_N^n,$$

where $\mu_N = e^{2\pi i/N}$ is the standard Nth root of unity.

For example, the Gauss sum of ψ above is zero, which does not tell us very much useful information.

Proposition 259

For any primitive character χ mod N and any integer m,

$$\sum_{n=0}^{N-1} \chi(n)\mu_N^{nm} = \overline{\chi}(m)g(\chi).$$

Notice the special cases when $m = N$ and $m = 1$: the sum of χ is zero for a primitive character.

Proof. When $\gcd(m, N) = 1$, this is simple: multiply the left side by $\overline{\chi}(m)\chi(m) = 1$ to get

$$\overline{\chi}(m)\sum_{n=0}^{N-1} \chi(n)\chi(m)\mu_N^{nm},$$

and now the sum is still summing over all integers mod N, so it is $g(\chi)$, recovering the desired result.

Otherwise, the right hand side is zero because $\overline{\chi}(m) = 0$. Let $d = \gcd(m, N)$, and set $m'd = m$ and $N'd = N$. Now rewrite the sum based on the residue of n mod N': for any given residue n' mod N', the exponents for those corresponding terms look like $\mu_N^{(n'+kN')m}$ for some k. But $d|m$, so the kN' terms disappear and we're just left with an exponent of $n'm$. Noting that $(\mu_N)^d = \mu_{N'}$, this leaves us with

$$\sum_{n=0}^{N-1} \chi(n)\mu_N^{nm} = \sum_{n'=0}^{N'-1}\left(\sum_{\substack{n=0\\ n\equiv n' \bmod N'}}^{N-1} \chi(n)\right)\mu_{N'}^{n'm'}.$$

To show the desired result, it suffices to show that the parenthetical term is zero. First of all, consider the sum over K (that is, when $n' = 1$): since χ is primitive, the value of χ over this kernel should not be trivial, so by a similar "sum-over-group" argument, the sum over K is zero. And now we can sum over any other coset $n'K$ and get zero by multiplicativity, and we're done. □

Corollary 260

The Gauss sum of a primitive character has magnitude $\sqrt{N}$ (in particular, it is nonzero).

Proof. We know that

$$g(\chi)\overline{g}(\chi) = g(\chi)\sum_{m=0}^{N-1} \overline{\chi}(m)\mu_N^{-m},$$

and bringing $g(\chi)$ in and using the above proposition yields

$$= \sum_{m=0}^{N-1}\sum_{n=0}^{N-1} \chi(n)\mu_N^{nm}\mu_N^{-m} = \sum_{n=0}^{N-1} \chi(n)\sum_{m=0}^{N-1} \mu_N^{(n-1)m},$$

where we've swapped the order of summation. But the sum over m (by a familiar argument) is zero unless $(n-1)m$ is a multiple of N, which happens exactly when $n = 1$. And that means the sum just reduces to $\sum_{n=0}^{N-1} 1 = N$, as desired. □

Proposition 261

If $N = 1$ or 2, then all Dirichlet characters have $\chi(-1) = 1$. Otherwise, there are an even number of Dirichlet characters, and half of them have each of $\chi(-1) = \pm 1$.

Proof. It's easy to write down the only Dirichlet character for $N = 1$ and 2: it's the trivial one. Otherwise, $\chi(-1)$ must either be -1 or 1 (because $\chi(-1)^2 = \chi(1) = 1$), and there does exist a character that takes -1 mod N to -1, because we can lift a nontrivial character either mod p for an odd prime $p|N$ (sending a primitive root mod p to a primitive root of unity) or the nontrivial character mod 4 when $4|N$ (which sends ± 1 to ± 1). So we have a surjective homomorphism from the group of characters $\hat{G}_N$ to $\{\pm 1\}$, and the correspondence theorem yields the result. □

This is most of what we'll need for the upcoming theory of Eisenstein spaces, and we'll finish with an important idea of decomposition: we'll construct subspaces of modular forms $\mathcal{M}_k(\Gamma_1(N))$ corresponding to the congruence subgroup of matrices $\gamma \equiv \begin{bmatrix} 1 & * \\ 0 & 1 \end{bmatrix}$ mod N.

Definition 262

The **χ-eigenspace** of a Dirichlet character χ mod N is

$$\mathcal{M}_k(N, \chi) = \{f \in \mathcal{M}_k(\Gamma_1(N)) : f[\gamma]_k = \chi(d_\gamma) f \quad \forall \gamma \in \Gamma_0(N)\},$$

where d_γ is the lower-right entry of γ.

This is clearly a (linear) subspace of $\mathcal{M}_k(\Gamma_1(N))$.

Lemma 263

$\mathcal{M}_k(N, \mathbf{1}) = \mathcal{M}_k(\Gamma_0(N))$, and $\mathcal{M}_k(N, \chi)$ is trivial unless $\chi(-1) = (-1)^k$.

Proof. For the first statement, note that $\mathbf{1}(d_\gamma) = 1$ because $\gcd(d_\gamma, N) = 1$, and the space of functions left invariant under $[\gamma]_k$ for $\gamma \in \Gamma_0(N)$ is exactly $\Gamma_0(N)$.

For the second statement, the statement is vacuously true for $N = 2$, and otherwise $-I$ is an element of $\Gamma_0(N)$. Then $f[\gamma]_k = (-1)^k f$, and the result follows. □

Proposition 264

The vector spaces $\mathcal{M}_k(\Gamma_1(N)), \mathcal{S}_k(\Gamma_1(N))$, and $\mathcal{E}_k(\Gamma_1(N))$ all decompose as direct sum of χ-eigenspaces: for example,

$$\mathcal{M}_k(\Gamma_1(N)) = \bigoplus_{\chi} \mathcal{M}_k(N, \chi),$$

and analogously for $\mathcal{S}_k$ and $\mathcal{E}_k$.

Proof. The proofs are the same for the first two spaces, and then quotienting yields the result for the third (since $(A \oplus B)/(C \oplus D) \cong A/C \oplus B/D$ for spaces $C \subset A, D \subset B$). We'll only present the first proof. Define the weight-k operator

$$\langle d \rangle = \left[\begin{bmatrix} a & b \\ c & \delta \end{bmatrix}\right]_k$$

for any $\delta \equiv d \bmod N$. This is a well-defined, multiplicative operator called a **diamond operator**, but we need more theory from Chapter 5 to see why this is indeed well-defined. Now for each character mod N, we can define

$$\pi_\chi = \frac{1}{\phi(N)} \sum_{d \in (\mathbb{Z}/n\mathbb{Z})^*} \chi(d)^{-1} \langle d \rangle.$$

This is a linear **projection operator**: applying π_χ twice as a double sum and using multiplicativity includes $\chi(d)$ for each element $\phi(N)$ times. To understand where this projection operator sends our modular forms, if we take any $f \in \mathcal{M}_k(\Gamma_1(N))$, then

$$\boxed{\pi_\chi f[\gamma]_k} = \frac{1}{\phi(N)} \sum_{d \in (\mathbb{Z}/n\mathbb{Z})^*} \chi(d)^{-1} \langle d \rangle f[\gamma]_k$$

still sums over d because the lower-right element d_γ of $[\gamma]$ is invertible:

$$= \frac{1}{\phi(N)} \sum_{d \in (\mathbb{Z}/n\mathbb{Z})^*} \chi(d_\gamma)\chi(d)^{-1}\chi(d_\gamma)^{-1} \langle dd_\gamma \rangle f = \boxed{\chi(d_\gamma)\pi_\chi f},$$

so we do project **into** $\mathcal{M}_k(N, \chi)$ ($\pi_\chi f$ satisfies the charactersitic condition), and when f is already in the χ-eigenspace, π_χ acts as

$$\pi_\chi f = \frac{1}{\phi(N)} \sum_d \chi(d)^{-1} \langle d \rangle f = \frac{1}{\phi(N)} \sum_d \chi(d)^{-1} \chi(d) f = f,$$

so we project **onto** the subspace.

To finish, note that the sum over characters

$$\sum_\chi \pi_\chi = \sum_\chi \frac{1}{\phi(N)} \sum_{d \in (\mathbb{Z}/n\mathbb{Z})^*} \chi(d)^{-1} \langle d \rangle$$

is zero except when $d = 1$, in which case we just end up with $\frac{1}{\phi(N)}\phi(N) \cdot \langle 1 \rangle$, which is the **identity operator** on Γ_0. This means the subspaces are spanning. Finally, for any two distinct characters χ, χ',

$$\pi_\chi \circ \pi_{\chi'} = \frac{1}{\phi(N)^2} \sum_{d,e \in (\mathbb{Z}/n\mathbb{Z})^*} \chi(d)^{-1} \chi'(e)^{-1} \langle de \rangle$$

by multiplicativity, but summing over any constant $c = de$ yields a sum

$$\sum_d \chi(d)^{-1} \chi'(d) \chi(c),$$

which is zero since $\chi^{-1}\chi$ is not the identity. Thus we've shown that the subspaces are disjoint: nothing projects down to both subspaces, and we're done. □

Diamond and Shurman 4.4 – Natalie Stewart

The basic structure of this lecture is to go through a laundry list of functions to define and use: we'll be discussing the gamma function Γ, zeta function ζ (and associated ξ), L-functions $L(s, \chi)$ and the modified zeta function $\zeta_+^n(s)$. The point is to provide meromorphic extensions to all of $\mathbb{C}$ for these functions, and characterize the poles and important explicit values.

Definition 265

The **gamma function**

$$\Gamma(s) = \int_0^\infty e^{-t} t^{s-1} dt.$$

This is an "extension of the factorial function:" we have that $\Gamma(n+1) = n!$ for integers n, and in fact whenever $\text{Re}(s) > 0$, the integral converges because we have exponential decay from e^{-t}. The main point is that we have a **functional equation** for all $\text{Re}(s) > 0$ of the form

$$\Gamma(s) = \frac{\Gamma(s+N)}{s(s+1)\cdots(s+n-1)},$$

which we prove via integration by parts of the identity $\Gamma(s+1) = s\Gamma(s)$ (we lower the power of the polynomial term). The point is that we can use this to extend Γ to a function on all of $\mathbb{C}$, except that we avoid points where we have to divide by zero:

$$\lim_{s \to -n} (s+n)\Gamma(s) = \lim_{s \to -n} \frac{\Gamma(s+n+1)}{s \cdot (s+1) \cdots (s+n-1)} = \frac{(-1)^n}{n!}.$$

So when we meromorphically continue Γ to the domain where $\text{Re}(s) \le 0$, we will have a simple pole at each integer $-n \in \mathbb{Z}_{\le 0}$ with a residue $\frac{(-1)^n}{n!}$. There's also another alternative formula that we can use:

Proposition 266

$\Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin(\pi s)}$.

In particular, this means that $\Gamma(s)$ is nonvanishing on all of $\mathbb{C}$ (we just need to check explicitly at the positive integers where we have poles, where we already know the value of Γ), so $\frac{1}{\Gamma}$ is an entire function.

Proof sketch. Break down both sides into series: we use the fact that

$$\lim \left(1 - \frac{t}{n}\right)^n = e^{-t}, \quad \sin(\pi s) = \pi s \prod_{m=1}^\infty \left(1 - \frac{s^2}{m^2}\right)$$

(the latter is the Euler product, which "tracks the zeros of the sine "). Then breaking down the integrals means that we want to show

$$I_n = \int_0^n \left(1 - \frac{t}{n}\right)^n t^{s-1} dt$$

converges to $\Gamma(s)$, and these are easier to integrate by parts because we are integrating polynomials. We find that

$$I_n(s) I_n(-s) = \frac{1}{-s^2 \prod_{m=1}^n (1 - s^2/m^2)},$$

and taking $n \to \infty$ will give us what we want. □

If we define the constant $C_k = \frac{(-2\pi i)^k}{\Gamma(n)}$, we can verify that $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$ (change variables to a Gaussian integral), and in general we can compute at half-integers that

$$\frac{\pi^{-(1-k)/2}\Gamma((1-k)/2)}{\pi^{-k/2}\Gamma(k/2)} = \frac{C_k}{2}.$$

This is shown by directly manipulating products:

$$\frac{C_k}{2} = \frac{(-2\pi i)^k}{2\Gamma(k)} = \frac{2^k \pi^k (-1)^{k/2}}{2(1 \cdot 3 \cdots (k-1)(2 \cdot 4 \cdots (k-2))},$$

and then we distribute factors of -2 into the parentheses on the left and $+2$ into the parentheses on the right. (Everything works out – the point is to set up a parity between powers of π and values where Γ is evaluated.)

We'll now move on to another function:

Definition 267

The **Riemann zeta function** is defined as

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

for any $\text{Re}(s) > 1$.

The relevant complex analysis argument shows that this is indeed absolutely convergent on this half-plane, and we also have the product formula

$$\zeta(s) = \prod_{p \text{ prime}} (1 - p^{-s})^{-1}$$

(we can basically multiply each factor for a prime p over to the other side, and that accounts for all of the denominators that have prime powers of p). We'll define in section 4.9 that if we define the function

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

we have the functional equation $\xi(s) = \xi(1-s)$, and the only poles of this function are at $s = 0, 1$. And this tells us that we have a meromorphic continuation of ζ to the entire complex plane, such that the only pole is at $s = 1$ (with residue 1) and we have simple zeros at $s = -2, -4, -6, \cdots$ (and possibly other points as well).

The main argument we'll make about these is that we can use our coefficients C_k from earlier to get all the negative integer values of ζ: recall that $\zeta(k) = \frac{C_k B_k}{2k}$, where B_k is the kth Bernoulli number, and that tells us that

$$\zeta(s) = \frac{\pi^{(1-s)/2}\Gamma((1-s)/2}{\pi^{-s/2}\Gamma(-s/2}\zeta(1-s),$$

so $\zeta(1-k) = \frac{2\zeta(k)}{C_k} = \frac{B_k}{k}$.

We can also generalize this zeta function using the Dirichlet character:

Definition 268

For a Dirichlet character χ mod N, define the **L-function**

$$L(s, \chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}.$$

for $\text{Re}(s) > 1$.

A similar product argument shows that

$$L(s, \chi) = \prod_p (1 - \chi(p)p^{-s})^{-1},$$

and this means that, for example,

$$L(s, \mathbf{1}_N) = \zeta(s) \prod_{p|N} (1 - p^{-s})$$

(remember that $\mathbf{1}_N$ is 1 whenever our integer is relatively prime to N and 0 otherwise). So at least for the identity character $\mathbf{1}_N$, we can meromorphically continue the L-function: $L(s, \mathbf{1}_N)$ will have a meromorphic continuation to the entire complex plane, where there is only a simple pole at 1 with residue $\prod_{p|N}(1 - p^{-s}) = \frac{\phi(N)}{N}$.

But whenever χ is not the character $\mathbf{1}_N$, we actually have better behavior: we can find functional equations for non-principal Dirichlet characters, and the extensions will turn out to be **entire**. So the key point is that ζ and $L(s, \chi)$ are meromorphic with some nice properties.

We'll quickly talk about the last family of functions here:

Definition 269

The **modified zeta functions** are defined via

$$\zeta_+^n(k) = \sum_{\substack{m=1 \\ m \equiv n \bmod N}}^{\infty} \frac{1}{m^k}, \quad \zeta^n(k) = \sum_{m \equiv n \bmod N}^{\prime} \frac{1}{m^k}.$$

(where we're allowed to sum over negative m in the second sum). Also, define

$$\zeta_+^n(k, \mu) = \sum_{\substack{m=1 \\ m \equiv n \bmod N}}^{\infty} \frac{\mu(m)}{m^k}.$$

We care about these because they showed up in our equations between our Eisenstein series $E_k^{\overline{v}}$ and $G_k^{\overline{v}}$, and in fact this gives us a relation between ζ_+^n and $\zeta_+^n(k, \mu)$. We can write our modified zeta functions as a sum of L-functions: we have that

$$\frac{1}{\phi(N)} \sum_{\chi \in \hat{G}_N} \chi(n^{-1}) L(s, \chi) = \zeta_+^n(s)$$

by using the orthogonality relations, and therefore ζ_+^n also has a meromorphic continuation to the complex plane, which again only has a simple pole at 1. This gives us the continuation of $\zeta_+^n(s, \mu)$ by defining linear relations between vectors of $E_k^{\overline{v}}$ and $G_k^{\overline{v}}$ and using a fair bit of linear algebra: the punchline is that we can write $\zeta_+^n(s, \mu)$ in terms of L-functions, so we can meromorphically continue it as well. Finally,

$$\zeta^n(s) = \zeta_+^n(s) + (-1)^{-s} \zeta_+^{-n}(s),$$

so we can find a meromorphic extension, and in fact ζ^n is **entire** (we just need to check that there isn't a pole at $s = 1$). One value that we will care about later on is

$$\zeta^n(1) = \frac{\pi i}{N} + \frac{\pi}{N} \cot \frac{\pi n}{N}.$$

21 April 30, 2020

Diamond and Shurman 4.5 – Nikhil Reddy

Recall that we defined the (original) Eisenstein series $G_k(\tau)$ and $E_k(\tau)$ by summing $\frac{1}{(c\tau+d)^k}$ over all integers (c, d) (except for zero), and we can also extract out a factor of $\zeta(k)$ by only considering coprime (c, d) to define the normalized Eisenstein series. We checked that both of these converge absolutely when $k \geq 3$, and we'll assume that throughout this section. (Notice that G_k and E_k are both identically zero in this definition whenever k is odd, so these are really only defined for even k.)

Our next idea was to define the alternate series

$$E_k^{\overline{v}}(\tau) = \varepsilon_N \sum_{\substack{(c,d)\equiv v \mod N \\ \gcd(c,d)+1}} \frac{1}{(c\tau+d)^k}, \quad G_k^{\overline{v}}(\tau) = \varepsilon_N \sum_{(c,d)\equiv v \mod N} \frac{1}{(c\tau+d)^k}$$

for a vector $\overline{v}$ of additive order N. (Remember that the overline means we reduce mod N.) These play nicely with each other as well, and notice that taking $N = 1$ gives us the original E_k and G_k.

We also proved that the $G_k^{\overline{v}}$s were a linear combination of the $E_k^{\overline{v'}}$s, given by

$$G_k^{\overline{v}} = \frac{1}{\varepsilon_N} \sum_{n \in (\mathbb{Z}/n\mathbb{Z})^*} \zeta_+^n(k) E_k^{\overline{n^{-1}v}}(\tau),$$

where $\zeta_+^n(k)$ is the modified Riemann zeta function. We'll see a return of the Dirichlet characters as well: these are the homomorphisms from $G_N \to \mathbb{C}^*$, and the important idea is that a character has a **conductor** (which is the periodicity as a function $\mathbb{Z} \to \mathbb{C}$): a primitive character is one with condutor N. (And remember that overline over a character is a conjugate, not a reduction mod N.) Recall that when χ is primitive, we have

$$\sum_{n=0}^{N-1} \chi(n)\mu_N^{nm} = \overline{\chi}(m)g(\chi)$$

and therefore the Gauss sum is nonzero for a primitive character. This is helpful for us to to decompose the space

$$\mathcal{M}_k(\Gamma_1(N)) = \bigoplus_{\chi} \mathcal{M}_k(N, \chi)$$

into χ-eigenspaces: we're going to be studying those eigenspaces in this lecture here.

We'll keep doing a bit more review so the calculations make sense: last week, we showed the weight-k operator satisfies

$$(E_k^{\overline{v}}[\gamma]_k)(\tau) = E_k^{\overline{v\gamma}}(\tau).$$

Since the $G_k^{\overline{v}}$s are linear combinations of multiple $E_k^{\overline{v'}}$s, **the same equation holds for them as well**:

$$(G_k^{\overline{v}}[\gamma]_k)(\tau) = G_k^{\overline{v\gamma}}(\tau).$$

we'll use this throughout the lecture.

Proposition 270

The Eisenstein series $G_k^{\overline{(0,d)}}$ for any $d \in (\mathbb{Z}/n\mathbb{Z})^*$ lives in $\mathcal{M}_k(\Gamma_1(N))$.

Proof. Any element $\gamma \in \Gamma_1(N)$ acts as the identity on vectors $(0, d)$, so the above equation tells us that

$$(G_k^{\overline{v}}[\gamma]_k)(\tau) = G_k^{\overline{v}}(\tau),$$

which is the characteristic property for a modular form in $\mathcal{M}_k(\Gamma_1(N))$. □

Our goal from here will be to construct a basis of the Eisenstein space by using pairs of primitive characters. Our first step will be to try to build an element of $\Gamma_0(N)$ from our elements $G_k^{\overline{(0,d)}}$:

Proposition 271

The Eisenstein series

$$\sum_{d\in(\mathbb{Z}/n\mathbb{Z})^*} G_k^{\overline{(0,d)}}$$

is an element of $\mathcal{M}_k(\Gamma_0(N))$.

Proof. Notice that applying the operator $[\gamma]_k$ means we're just summing over the group in a different order (using dd_γ instead of d, so we have the same sum again when we apply that operator. □

Proposition 272

The Eisenstein series

$$\sum_{d\in(\mathbb{Z}/n\mathbb{Z})^*} \overline{\chi}(d)G_k^{\overline{(0,d)}}$$

is an element of the χ-eigenspace $\mathcal{M}_k(N,\chi)$.

Proof. We're still bringing the d_γ into the vector, so the result under the weight-k operator is

$$\sum_{d\in(\mathbb{Z}/n\mathbb{Z})^*} \overline{\chi}(d)G_k^{(0,dd_\gamma)}.$$

Multiplying in a factor of $\chi(d_\gamma)\overline{\chi(d_\gamma)} = 1$ gives us a $\chi(d_\gamma)$ that comes out, and the rest is "summing over the group" again. Then we end up with

$$\chi(d_\gamma) sum_{d\in(\mathbb{Z}/n\mathbb{Z})^*}\overline{\chi}(d)G_k^{(0,dd_\gamma)},$$

which is the definition of the χ-eigenspace. □

Now we can do the full construction as follows: start with two Dirichlet characters ψ mod u and ϕ mod v, such that

- $uv = N$,
- The product character satisfies $(\psi\chi)(-1) = (-1)^k$, and
- ϕ is primitive.

Then the definition we're using is

$$G_k^{\psi,\phi}(\tau) = \sum_{c=0}^{u-1}\sum_{d=0}^{v-1}\sum_{e=0}^{u-1}\psi(c)\overline{\phi(d)}G_k^{\overline{(cv,d+ev)}}(\tau).$$

This is secretly pretty familiar: if we plug in $u = 1$, the sums over c and e disappear, and we're just left with the series we just worked with in the ϕ-eigenspace in the proposition above. (And we define $G_k^{\overline{v}}$ to be zero when $\overline{v}$ is not of order N.)

Let's prove this is in the ϕ-eigenspace: notice that

$$\overline{(cv, d+ev)\gamma} = \overline{(c'v, d'+e'v}$$

where we change the coordinates via $c' = ca_\gamma, d' = dd_\gamma, e' = (e + c'b_\gamma)d_\gamma$, so we might as well sum over c', d', e'

instead. So the action under $[\gamma]_k$ yields

$$G_k^{\psi,\phi}(\tau)[\gamma]_k = \sum_{c'=0}^{u-1}\sum_{d'=0}^{v-1}\sum_{e'=0}^{u-1} \psi(c)\overline{\phi(d)}G_k^{\overline{(c'v,d'+e'v)}}(\tau).$$

and now we just need to change the c and d in the Dirichlet characters: we use the fact that

$$\psi(c)\overline{\phi(d)} = \overline{\psi(a_\gamma)}\psi(c')\phi(d_\gamma)\overline{\phi}(d'),$$

and becaus $a_\gamma d_\gamma = 1$, we can write $\overline{\psi(a_\gamma)} = \psi(d_\gamma)$, so this just means we end up with

$$G_k^{\psi,\phi}[\gamma]_k = (\psi\phi)(d_\gamma)G_k^{\psi,\phi}.$$

This gives us our first result:

Theorem 273

The Eisenstein series $G_k^{\psi,\phi}$ lives in $\mathcal{M}_k(N,\psi\phi)$.

Remember that we have a Fourier series for $G_k^{\overline{v}}$, where the non-constant part looks like

$$\frac{C_k}{N^k}\sum_{mn>0,n\equiv c_v \bmod N} \operatorname{sgn}(m)m^{k-1}\mu_N^{d_v m}q_N^{nm}.$$

We can plug this into the formula for our Eisenstein series $G_k^{\psi,\phi}$ to get a quadruple sum

$$= \frac{C_k}{N^k}\sum_{c=0}^{u-1}\sum_{d=0}^{v-1}\sum_{e=0}^{u-1}\psi(c)\overline{\phi}(d)\sum_{mn>0,n\equiv c_v} \operatorname{sgn}(m)m^{k-1}\mu_N^{(d+ev)m}q_N^{mn}$$

which we can group by summing roots of unity, using that $uv = N$:

$$= \frac{C_k}{N^k}\sum_{c=0}^{u-1}\sum_{d=0}^{v-1}\psi(c)\overline{\phi}(d)\sum_{mn>0,n\equiv c_v} \operatorname{sgn}(m)m^{k-1}\mu_N^{dm}\left(\sum_{e=0}^{u-1}\mu_u^{em}\right)q_N^{mn}$$

But the inner sum of roots of unity is zero unless u divides m, and also n is a multiple of v by definition. So we can get rid of one of the summations by replacing $m \to um$ and $n \to vn$ to get

$$= \frac{C_k}{v^k}\sum_{c=0}^{u-1}\sum_{d=0}^{v-1}\psi(c)\overline{\phi}(d)\sum_{mn>0,n\equiv c \bmod u} \operatorname{sgn}(m)m^{k-1}\mu_v^{dm}q^{mn},$$

and now because ϕ is primitive by assumption, we can bring the $\overline{\phi}(d)$ into the inner sum and use our proposition about Gauss sums: now we have

$$\frac{C_k g(\overline{\phi})}{v^k}\sum_{c=0}^{u-1}\psi(c)\sum_{mn>0,n\equiv c(u)} \operatorname{sgn}(m)\phi(m)m^{k-1}q^{mn},$$

and now the sum c can be summed over all n instead to get

$$= \frac{C_k g(\overline{\phi})}{v^k}\sum_{mn>0}\sum_{mn>0} \operatorname{sgn}(m)\psi(n)\phi(m)m^{k-1}q^{mn},$$

and now the negativity condition tells us that the (m,n) and $(-m,-n)$ terms are identical: we end up with

$$= \frac{C_k g(\overline{\phi}}{v^k}\sum_{m,n>0}\psi(n)\phi(m)m^{k-1}q^{mn},$$

and we can rewrite this in a way that looks like the σ_{k-1} coefficients in the original Eisenstein series:

$$= \frac{C_k g(\overline{\phi})}{v^k} 2 \sum_{n=1}^{\infty} \left(\sum_{m|n, m>0} \psi(n/m)\phi(m)m^{k-1} \right) q^m.$$

We can then derive the following formula:

Theorem 274

We can rewrite the Eisenstein series as

$$G_k^{\psi,\phi} = \frac{C_k g(\overline{\phi})}{v^k} E_k^{\psi,\phi},$$

where

$$E_k^{\psi,\phi} = \delta(\psi)L(1-k,\phi) + 2\sum_{n=1}^{\infty} \sigma_{k-1}^{\psi,\phi}(n)q^n,$$

where $\delta(\psi)$ is 1 when the character is identically 1 and 0 otherwise, and we have the generalized power sum

$$\sigma_{k-1}^{\psi,\phi} = \sum_{m|n, m>0} \psi(n/m)\phi(m)m^{k-1}$$

So now we'll move to constructing the basis elements:

Definition 275

Let $A_{N,k}$ be the set of triples (ψ, ϕ, t) such that ψ, ϕ are primitive mod u and v with $(\psi\phi)(-1) = (-1)^k$, and t is an integer such that $tuv|N$. Let $B_{N,k}$ be the set of pairs (ψ', χ') such that ψ' and χ' are (not necessarily primitive) characters modulo u', v' with $u'v' = N$ and $(\psi'\phi')(-1) = (-1)^k$.

These sets are **in bijection**, because we can send

$$(\psi, \phi, t) \to (\psi'_{tu}, \phi'_{N/tu})$$

and

$$(\psi', \phi') \to (\psi, \phi, u'/u)$$

where ψ, ϕ are the primitive characters mod u' and v' (the v'/v is implied because we can determine it from the value of N). This is important because $B_{N,k}$ is easier to work with – we don't need to worry about primitive characters – at first-order, the size of $B_{N,k}$ is

$$\frac{1}{2}\sum_{d|N} \phi(d)\phi(N/d),$$

and this seems to be also the size of $\mathcal{E}_k(\Gamma_1(N))$. This turns out to be true for all $N \geq 4$, and thus it makes sense that we can use the set $A_{N,k}$ to make a basis (using our $G_k^{\psi,\phi}$ functions):

Theorem 276

Let N be a positive integer and $k \geq 3$. Define $E_k^{\psi,\phi,t}(\tau) = E_k^{\psi,\phi}(t\tau)$: then the set $E_k^{\psi,\phi,t}$ form a basis of $\mathcal{E}_k(\Gamma_1(N))$. In particular, the set of $E_k^{\psi,\phi,t}$ with $\psi\phi = \chi$ also give us a basis for the χ-eigenspace $\mathcal{E}_k(N, \chi)$.

In particular, choosing $\psi = 1$ also gives us a basis of eigenspaces for Γ_0.

Diamond and Shurman 4.6 – Christian Altamirano

In this section, we'll talk about the Eisenstein spaces for $k = 2$. We'll begin by constructing a basis for $\mathcal{E}_2(\Gamma(N))$: the Weierstrass $\wp$ function

$$\wp(z) = \frac{1}{z^2} + \sum_{\omega \in \Lambda}{}' \frac{1}{(z-\omega)^2} - \frac{1}{\omega^2}$$

helps us define the function (where we use the lattice $\mathbb{Z} + \mathbb{Z}\tau$)

$$f_2^{\overline{v}} = \frac{1}{N^2}\wp_\tau\left(\frac{c_v(\tau + d_v}{N}\right)$$

which is a weakly modular function of weight 2 with respect to $\Gamma(N)$. We'll try and expand that now: this yields

$$f_2^{\overline{v}} = \frac{1}{(c_v\tau + d_v)^2} + \frac{1}{N^2}\sum_{(c,d)\in\mathbb{Z}^2}{}' \frac{1}{\left(\frac{c_v\tau+d_v}{N} - c\tau - d\right)^2} - \frac{1}{(c\tau + d)^2}.$$

Our goal is to compute the Fourier coefficients here. One formula that we'll use is the Fourier series of the following sum:

Proposition 277

For any $\tau \in \mathcal{H}$ and $k \geq 2$, we have

$$\sum_{d\in\mathbb{Z}} \frac{1}{(\tau + d)^k} = C_k \sum_{m=1}^{\infty} m^{k-1} q^m,$$

and similarly if $-\tau \in \mathcal{H}$, we have

$$\sum_{d\in\mathbb{Z}} \frac{1}{(\tau + d)^k} = -C_k \sum_{m=-\infty}^{-1} m^{k-1} q^m.$$

We'll use the same notation as in section 4.2, since there will be a lot of computation.

Theorem 278

We have

$$f_2^{\overline{v}}(\tau) = G_2^{\overline{v}}(\tau) - \frac{1}{N^2}G_2(\tau),$$

where $G_2(\tau)$ is the regular Eisenstein series of weight 2 with Fourier expansion $2\zeta(2) + 2C_2\sum_{n=1}^{\infty}\sigma(n)q^n$ and

$$G_2^{\overline{v}}(\tau) = \delta(c_v)\zeta^{d_v}(2) + \frac{C_2}{N^2}\sum_{n=1}^{\infty}\sigma_1^{\overline{v}}(n)q_N^n.$$

Proof. We'll split the sum for $f_2^{\overline{v}}(\tau)$ into three parts: $c = 0, c > 0$, and $c < 0$. When $c = 0$, we'll also add the constant term $\frac{1}{(c_v\tau+d_v)^2}$. Then we end up with a sum

$$\frac{1}{N^2}\sum_{d\in\mathbb{Z}}\left(\frac{1}{\left(\frac{c_v\tau+d_v}{N} - d\right)^2}\right) - \frac{1}{N^2}\sum_{d\in\mathbb{Z}}{}' \frac{1}{d^2}.$$

The right sum is just $\frac{1}{N^2}2\zeta(2)$, and the left sum depends on c_v: if $c_v = 0$, we just get $\zeta^{d_v}(2)$, and otherwise the imaginary part of $c_v\tau + d_v$ is positive (because we only sum c_v from 0 to N), so we can use the Fourier expansion in the proposition above using $c_v\tau + d_v$ instead of τ. This yields

$$\frac{1}{N^2} C_2 \sum_{m=1}^{\infty} m^{k-1} (q')^m,$$

where $q' = e^{2\pi i (c_v\tau + d_v)/N}$. Then q' is simply related to q, so we can plug that in: this means that $c = 0$ contribution is just

$$\delta(c_v)\zeta^{d_v}(2) + (1 - \delta(c_v)) \frac{C_2}{N^2} \sum_{m=1}^{\infty} \mu_N^{d_v m} q_N^{c_v m} - \frac{1}{N^2} 2\zeta(2).$$

For the $c > 0$ case, we want to figure out

$$\frac{1}{N^2} \sum_{c>0} \sum_{d \in \mathbb{Z}} \frac{1}{\left(\frac{c_v\tau + d_v}{N} - c\tau - d\right)^2} - \frac{1}{(c\tau + d)^2}.$$

We can apply the proposition for the second term again because $\operatorname{Im}(c\tau) > 0$, and we end up with

$$\frac{1}{N^2} \sum_{c>0} C_2 m (q')^m,$$

where $q' = e^{2\pi i c\tau} = q^c$. Reordering the sum by summing over mc instead, this becomes

$$= \frac{C_2}{N^2} \sum_{n=1}^{\infty} \sigma(n) q^n.$$

For the first term of this sum, we have a negative imaginary part, so we expand out the expressions there as well. The eventual result is that the $c > 0$ contribution looks like

$$\frac{C_2}{N^2} \sum_{n=1}^{\infty} \left(\sum_{\substack{m<0 \\ m|n \\ n/m \equiv c_v (N)}} \operatorname{sgn}(m) \mu_N^{d_v m} \right) q_N^n - \frac{C_2}{N^2} \sum_{n=1}^{\infty} \sigma(n) q^n.$$

And the $c < 0$ case is similar – the right sum is the same and we need to be careful about the negative sign for $\operatorname{Im}(\tau)$. Summing everything together, we do end up with the result that we wanted – for instance, things cancel out between $c = 0$ and $c < 0$. □

We'll now move on to the "correction terms" for these series: remember that the Eisenstein series $G_2(\tau)$ is only invariant when we add the correction term $-\frac{\pi}{\operatorname{Im}(\tau)}$. So something similar occurs here when we correct our G_2 expression, because

$$f_2^{\overline{v}}(\tau) = G_2^{\overline{v}}(\tau) - \frac{1}{N^2} G_2(\tau)$$

is weight-2 invariant.

Definition 279

Define the weight-2 modular form

$$g_2^{\overline{v}}(\tau) = G_2^{\overline{v}}(\tau) - \frac{\pi}{N^2 \operatorname{Im}(\tau)}.$$

To show holomorphicity everywhere, we can instead use the alternative method of proving the coefficients to be polynomially bounded:

Theorem 280

The coefficients of $g_2^{\overline{v}}$ are bounded by Cn^2.

Proof. We can make an easy bound: all coefficients just look like

$$|\sigma_1^{\overline{v}}(n)| \le \sum_{m|n, n/m \equiv c_v \mod N} |\text{sgn}(m) m \mu_N^{d_v m}| \le \sum_{m \le n} |m| \le Cn^2.$$

□

So we do indeed have a modular form in our above definition.

With this, we'll move on to calculating a basis for $\mathcal{E}_2(\Gamma(N))$ and $\mathcal{E}_2(\Gamma_1(N))$: remember that the dimension of $k = 2$ is $\varepsilon_\infty - 1$, so we can get a basis by taking the cusp representatives and taking differences to remove one basis element. It turns out that $\mathcal{E}_2(\Gamma(N))$ has basis

$$\{g_2^{\overline{v_1}} - g_2^{\overline{v_2}}, \cdots, g_2^{\overline{v_{\varepsilon_\infty} - 1}} - g_2^{\overline{v_{\{\varepsilon_\infty}}}\}.$$

We can also change these differences so that we use linear combinations over $G_2^{\overline{v}}$ instead of $g_2^{\overline{v}}$, as long as the corrections match up:

Theorem 281

The space $\mathcal{E}_2(\Gamma(N))$ can also be written as

$$\mathcal{E}_2(\Gamma(N)) = \left\{ \sum_{\overline{v}} a_{\overline{v}} G_2^{\overline{v}} : \sum_{\overline{v}} a_{\overline{v}} = 0 \right\}.$$

For $\mathcal{E}_2(\Gamma_1(N))$, we'll define the Eisenstein series $G_2^{\psi,\phi}(\tau)$ and $E_2^{\psi,\phi}(\tau)$ analogously to how we did in the previous lecture for $G_k^{\psi,\phi}(\tau)$ and $E_k^{\psi,\phi}(\tau)$.

Whenever ψ or ϕ is nontrivial here, $G_2^{\psi,\phi}$'s coefficients sum to zero, and also note that $G_2^{\psi,\phi} \in \mathcal{M}_2(N, \psi\phi)$ and that

$$G_2^{\psi,\phi}(\tau) = \frac{C_2 g(\phi)}{v^2} E_2^{\psi,\phi}(\tau).$$

But when ψ, ϕ are both trivial, we won't get a modular form because of the correction term. But we can still use that case to get a modular form:

Lemma 282

For an integer t, the function

$$C_2(E_2^{1_1,1_1}(\tau) - tE_2^{1_1,1_1}(t\tau)) = G_{2,t}(\tau) = G_2(\tau) - tG_2(t\tau)$$

is a modular form.

Proof. It suffices to show that

$$C_2 E_2^{1_1,1_1}(\tau) = G_2(\tau)$$

for any τ, and this is true because plugging in the trivial characters into the definition of E_2 yields

$$L(-1, 1_1) + 2\sum_{n=1}^{\infty} \sigma(n) q^n = \zeta(-1) + 2\sum_{n=1}^{\infty} \sigma(n) q^n,$$

but $G_2(\tau)$ is also of this form (we just need to check that $2\zeta(2) = C_1\zeta(-1)$ so that the constants match up). □

This means that we can finally define a basis for $\mathcal{E}_2(\Gamma_1(N))$: much like the last section, let $A_{N,2}$ be the set of triples (ψ, ϕ, t) such that ψ, ϕ are primitive characters mod u and v with $(\psi\phi)(-1) = 1$, and let t be an integer such that $1 < tuv|N$. (We don't want $t = u = v = 1$ here, so we're also excluding the tuple corresponding to the correction in the above lemma).

Definition 283

For any $(\psi, \phi, t) \in A_{N,2}$, define

$$E_2^{\psi,\phi,t}(\tau) = \begin{cases} E_2^{1_1,1_1}(\tau) - tE_2^{1_1,1_1}(t\tau) & \psi = \phi = 1_1 \\ E_2^{\psi,\phi}(t\tau) & \text{otherwise.} \end{cases}$$

And this set turns out to be the basis for $\mathcal{E}_2(\Gamma_1(N))$: in fact, the subsets

$$\{E_2^{\psi,\phi,t} : (\psi, \phi, t) \in A_{N,2}, \psi\phi = \chi\}$$

will form bases for the χ-eigenspaces $\mathcal{E}_2(N, \chi)$.

22 May 5, 2020

Diamond and Shurman 4.7 – Michael Tang

We'll be discussing the Bernoulli numbers and Hurwitz zeta function – this is mostly preparation for the next lecture. This will give a few formulas that are important about weight-1 Eisenstein series.

As a roadmpa, the first part of this talk will be about Bernoulli numbers (derived from the context of power sums) Bernoulli polynomials, and the Bernoulli numbers of a Dirichlet character. Then in the second part, we will do some sketchy complex analysis to generalize both ζ and the modified ζ function from previous lectures, and we'll relate them to Bernoulli numbers with a formula.

Definition 284

The ***k*th power sum** is defined as

$$S_k(n) = \sum_{m=0}^{n-1} m^k = 0^k + 1^k + \cdots + (n-1)^k,$$

where we define $0^0 = 1$ and $0^k = 0$ otherwise.

We're zero-indexing here so that computation later will be a bit more convenient.

Definition 285

The **generating function for the *k*th power sum** is the exponential generating function

$$\mathbb{S}(n, t) = \sum_{k=0}^{\infty} S_k(n) \frac{t^k}{k!}.$$

(Here, we should think of t as our variable and n as a parameter.)

Lemma 286

We have

$$\mathbb{S}(n,t) = \frac{e^{nt}-1}{e^t-1}.$$

Proof. We have

$$\mathbb{S}(n,t) = \sum_{k=0}^{\infty} S_k(n)\frac{t^k}{k!} = \mathbb{S}(n,t) = \sum_{k=0}^{\infty}\sum_{m=0}^{n-1} m^k \frac{t^k}{k!},$$

and swapping the order of summation turns this into

$$\sum_{m=0}^{n-1}\sum_{k=0}^{\infty}\frac{(mt)^k}{k!},$$

and now the inner sum is the Taylor series for e^{mt}. Evaluating the subsequent geometric series yields the result. □

We'll split this fraction up as

$$\mathbb{S}(n,t) = \frac{e^{nt}-1}{t}\frac{t}{e^t-1}.$$

This is because both functions here have a limit as we approach 0, and we can write both of these more easily.

Definition 287

The Bernoulli numbers are the coefficients of the exponential generating function

$$\frac{t}{e^t-1} = \sum_{k=0}^{\infty} B_k\frac{t^k}{k!}.$$

Substituting this in,

$$\mathbb{S}(n,t) = \frac{e^{nt}-1}{t}\sum_{j=0}^{\infty} B_j\frac{t^j}{j!},$$

and now we substitute in the Taylor series for e^{nt}, reparameterize the exponent of t after simplification, and we end up with

$$= \sum_{k=0}^{\infty}\left(\frac{1}{k+1}\sum_{j=0}^{k}\binom{k+1}{j}B_j n^{k+1-j}\right)\frac{t^k}{k!}.$$

But remember that $\mathbb{S}(n,t)$ is a generating function, so the coefficients must match up: this expression involving Bernoulli numbers must give us the power sum $S_k(n)$. We'll take a part of this expression:

Definition 288

The **Bernoulli polynomials** are defined as

$$B_k(X) = \sum_{j=0}^{k}\binom{k}{j}B_j X^{k-j}.$$

(Notation-wise, B_k denotes the number, and $B_K(X)$ denotes the polynomial.) So we can rewrite

$$S_k(n) = \frac{1}{k+1}\sum_{j=0}^{k}\binom{k+1}{j}B_j n^{k+1-j}, \quad B_k(X) = \frac{1}{k+1}(B_{k+1}(n) - B_{k+1})$$

(where we need to subtract off the last term of the sum in the definition of the Bernoulli polynomial because of the limits).

Example 289

We can look at the case $k = 2$.

Then we know that $B_0 = 1, B_1 = -\frac{1}{2}, B_2 = \frac{1}{6}$, and $B_3 = 0$ by looking at the Taylor expansion of $\frac{t}{e^t - 1}$, so

$$B_3(X) - B_3 = B_0 X^3 + 3B_1 X^2 + 3B_2(X) = X^3 - \frac{3}{2}X^2 + \frac{1}{2}X.$$

Substituting in $X = n + 1$ means that we have a formula for the sum of the first n squares – this naturally generalizes to finding a formula for the sum of kth powers.

Definition 290

Let $\psi : \mathbb{Z}/u\mathbb{Z}$ be a function (which doesn't need to be a Dirichlet character, but that's the case we care about). The Bernoulli numbers of ψ, $B_{k,\psi}$, are the constants such that

$$\sum_{k=0}^{\infty} B_{k,\psi} \frac{t^k}{k!} = \sum_{c=0}^{u-1} \psi(c) \frac{te^{ct}}{e^{ut} - 1}.$$

This may seem a bit unintuitive, but it has nice properties. If we plug in the trivial character $\chi = \mathbf{1}$, then $B_{k,\psi} = B_k$ because the right hand side simplifies to $\frac{t}{e^t - 1}$. And in fact, we can solve for $B_{k,\psi}$ in terms of the B_k:

$$B_{k,\psi} = u^{k-1} \sum_{c=0}^{u-1} \psi(c) B_k \left(\frac{c}{u} \right).$$

This is proved by swapping the order of summation again (plug in the generating function for the Bernoulli polynomials $B_k(c/u)$ on the right-hand side and do some more algebra). We'll need the $k = 1$ case in the next lecture, which says that

$$B_{1,\psi} = \sum_{c=0}^{u-1} \psi(c) \left(\frac{c}{u} - \frac{1}{2} \right).$$

Let's move on:

Definition 291

Let $r \in (0, 1]$. The **Hurwitz zeta function** is defined for all $\mathrm{Re}(s) > 1$ by

$$\zeta(s, r) = \sum_{n=0}^{\infty} \frac{1}{(r + n)^s}.$$

This converges absolutely for $\mathrm{Re}(s) > 1$ by the same argument as for ζ. Notice that when we plug in $r = 1$, this yields the usual Riemann zeta function (notice that the Hurwitz zeta function starts the sum from 0 instead of 1), and when we plug in a rational number $r = \frac{d}{N}$, we have $\zeta(s, r) = N^s \zeta_+^d(s)$ (both of these can be directly checked).

Theorem 292

Let $\psi \neq \mathbf{1}$ be a Dirichlet character mod u. Then the L-series

$$L(1-k,\psi) = \sum_{n=1}^{\infty} \frac{\psi(n)}{n^{1-k}} = -\frac{B_{k,\psi}}{k}.$$

In other words, we will be able to evaluate the L-function for Dirichlet characters at non-positive integers.

Proof sketch. First, we can relate the L-function to the Hurwitz function:

$$\sum_{c=1}^{u} \psi(c)\zeta\left(1-k, \frac{c}{u}\right) = u^{1-k} \sum_{c=1}^{u} \psi(c)\zeta_{+}^{c}(1-k) = u^{1-k} L(1-k,\psi),$$

where we've used the formula for $\zeta(s,r)$. We'll work more with the left hand side here: lots of complex analysis gives us the analytic continuation

$$\zeta(s,r) = -\frac{\Gamma(1-s)}{2\pi i} \int_{\gamma_\varepsilon} \frac{ze^{rz}}{e^z-1} z^{s-1} \frac{dz}{z},$$

for all $s \in \mathbb{C}$, where γ_ε (for $\varepsilon \in \mathbb{R}$) is a contour which goes from $-\infty$ to $-\varepsilon$, does a counterclockwise circle, and goes back to $-\infty$. Showing this requires the use of the Mellin transform, which we'll talk about later on.

Substituting this into our above formula yields

$$u^{1-k} L(1-k,\psi) = -\frac{\Gamma(k)}{2\pi i} \int_{\gamma_\varepsilon} \sum_{c=1}^{u} \psi(c) \frac{(c/u)e^{rc/u}}{e^{c/u}-1} \frac{dz}{z^{k+1}}.$$

Then if we apply the Cauchy integral formula to the integral around the circle, we indeed end up with the desired result (the integral along the line segments in γ_ε cancel out). □

So we now have a formula for our L-series on the nonnegative integers, and again we'll care about the $k=1$ case: this reduces to

$$L(0,\psi) = -B_{1,\psi}.$$

What's interesting here is that the series $L(0,\psi)$ is technically supposed to diverge, but we assign a value using this analytic continuation.

Diamond and Shurman 4.8 – Zack Chroman

The goal of this section is to continue to find concrete bases, but this time for the weight-1 eigenspaces $\mathcal{E}_1(N,\chi)$ and $\mathcal{E}_1(\Gamma_1(N))$. We'll mimic the $k \geq 3$ case as much as possible, but this is the most ugly case. Our approach will be to define a function that's almost the Weierstrass $\wp$ function, which will help us define a weight-1 modular form g_1. This will then extend to a function G_1 which will allow us to define the basis elements $G_1^{\psi,\phi}$ and $E_1^{\psi,\phi}$.

Definition 293

The **Weierstrass zeta function** is defined as

$$Z_\Lambda(z) = \frac{1}{z} + \sum_{\omega\in\Lambda}{}' \left(\frac{1}{z-\omega} + \frac{1}{\omega} + \frac{z}{\omega^2}\right).$$

The proof that this converges is similar to the proof that the Weierstrass function converges, and it turns out that we have

$$Z'_\lambda(z) = -\wp(z).$$

This function will not be doubly periodic like $\wp$ is, but its derivative is doubly periodic (as already established), so we get the lattice constants

$$\eta_1(\Lambda) = Z_\lambda(z+\omega_1) - Z_\Lambda(z), \quad \eta_2(\Lambda) = Z_\lambda(z+\omega_2) - Z_\Lambda(z)$$

(because their derivative is 0 – this is the same argument as usual).

Lemma 294

Let ω_1, ω_2 generate our lattice Λ. Then we have

$$\eta_2(\Lambda)\omega_1 - \eta_1(\Lambda)\omega_2 = 2\pi i$$

when $\frac{\omega_1}{\omega_2} \in \mathbb{H}$.

Proof. The proof mimics the way we understand the Weierstrass $\wp$ function: we have a fundamental parallelogram with boundary ∂P, and we can translate this by t so that there are no poles on the boundary. The residue formula tells us that

$$\int_{t+\partial P} Z_\Lambda(z)dz = 2\pi i$$

because the residues are 1, but we can also break this up into parts as

$$\int_0^{\omega_1} (Z_\Lambda(z+t) - Z_\Lambda(z+\omega_2+t))dz + \int_0^{\omega_2} (-Z_\Lambda(z+t) + Z_\Lambda(z+\omega_1+t))dz$$

(one term for each side of the parallelogram), and this is exactly $\eta_2(\Lambda)\omega_1 - \eta_1(\Lambda)\omega_2$. □

We're going to black box some facts: it turns out that when Λ is generated by 1 and τ, we have

$$Z_{\Lambda_\tau}(z) = \eta_2(\Lambda_\tau)z - \pi i \frac{1+e^{2\pi i z}}{1-e^{2\pi i z}} = -2\pi i \sum_{n=1}^{\infty} \left(\frac{e^{2\pi i z q^n}}{1-e^{2\pi i z q^n}} - \frac{e^{-2\pi i z q^n}}{1-e^{-2\pi i z q^n}} \right)$$

where $q = e^{2\pi i \tau}$. We also have that $\eta_2(\Lambda_\tau) = G_\tau)$, and therefore $\eta_1(\Lambda_\tau) = \tau G_2(\tau) - 2\pi i$ by the previous lemma.

Now we can define our functions g_1: let Λ be generated by ω_1 and ω_2. Fix an integer N and a vector $(c_v, d_V) \in (\mathbb{Z}/N\mathbb{Z})^2$, and we define

$$F_1^{\overline{v}}(\Lambda) = Z_\Lambda\left(\frac{c_v\omega_1 + d_v\omega_2}{N}\right) - \frac{c_v\eta_1(\Lambda) - d_v\eta_2(\Lambda)}{N}.$$

We can check that the contributions cancel out if we add N to c_v or d_v, so this is well-defined. When we specialize to the case where Λ is generated by 1 and τ, we also define

$$g_1^{\overline{v}}(\tau) = \frac{1}{N} Z_{\Lambda_\tau}\left(\frac{c_v\tau + d_v}{N}\right) - \frac{c_v\eta_1(\Lambda_\tau) - d_2\eta_2(\Lambda_\tau)}{N^2}.$$

Lemma 295

g_1 is weight-1 invariant.

We'll also black box this fact. Our short-term goal with this is to **plug in the formula** for $Z_{\Lambda_\tau}(z)$ into this formula here for g_1 to extract more information. We'll define $z = \frac{c_v\tau+d_v}{N}$, $q = e^{2\pi i\tau}$, $q_N = e^{2\pi i\tau/N}$, and $\mu_N = e^{2\pi i/N}$ as we have in the past few lectures.

We'll go step-by-step from here when we plug in the complicated formula for Z_{Λ_τ}. It turns out that the first and last terms cancel out nicely:

$$\frac{\eta_2(\Lambda_\tau)(z)}{N} - \frac{c_v\eta_1(\Lambda_\tau) - d_v\eta_2(\Lambda_\tau)}{N^2} = \frac{2\pi i c_v}{N^2}.$$

We'll now deal with the

$$-\frac{\pi i}{N}\frac{1+e^{2\pi iz}}{1-e^{2\pi iz}}$$

term. In the case where $c_v = 0$, then we just have $z = d_v$, and this simplifies to $\frac{\pi}{N}\cot\frac{\pi d_v}{N}$. Otherwise, we'll expand as a geometric series to get

$$\frac{1+e^{2\pi iz}}{1-e^{2piiz}} = 1 + 2\sum_{m=1}^{\infty} e^{2\pi im(c_v\tau+d_v)/N} = 1 + 2\sum_{m=1}^{\infty} q_N^{c_v m}\mu_N^{d_v m},$$

meaning we can just write

$$-\frac{\pi i}{N}\frac{1+e^{2\pi iz}}{1-e^{2\pi iz}} = \delta(c_v)\frac{\pi}{N}\cot\frac{\pi d_v}{N} + (1-\delta(c_v))\left(-\frac{\pi i}{N} + \frac{C_1}{N}\sum_{m=1}^{\infty} q_N^{c_v m}\mu_N^{d_v m}\right)$$

where $\delta(c_v)$ is the indicator function for $\overline{c_v}$ being zero, and $C_1 = -2\pi i$ (to draw an analogy with the previous cases).

Finally, we need to deal with the other sums. First, we deal with

$$\sum_{n=1}^{\infty}\frac{e^{-2\pi izq^n}}{1-e^{-2\pi izq^n}}.$$

We can expand out the geometric series, replace q with q_N^N, and replace n with $(c_v\tau + d_v)/N$, and we rearrange in the same way as the last sum to get to

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} e^{-2\pi imd_v/N} q_N^{nmN-mc_v}.$$

The q_N term is the important part, so we can reparameterize as

$$= \sum_{k=1}^{\infty}\sum_{\substack{\ell<0 \\ \ell|k \\ k/\ell\equiv c_v}} \mu_N^{md_v} q_N^k$$

where $k = nmN - mc_v$ and $\ell = -m$.

Now, we need to deal with

$$\sum_{n=1}^{\infty}\frac{e^{2\pi izq^n}}{1-e^{2\pi izq^n}},$$

and we get to the same kind of intermediate statement

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} e^{2\pi imd_v/N} q_N^{nmN+mc_v}.$$

But this time, we overcount some extra terms, so we need to correct that: it only happens for the $n = 0$ term, which doesn't appear in our original sum. This yields

$$= \left(\sum_{k=1}^{\infty} \sum_{\substack{m>0 \\ m|k \\ k/m \equiv c_v}} \mu_N^{md_v} q_N^k \right) - (1-\delta(c_v)) \sum_{m=1}^{\infty} \mu_N^{md_v} q_N^{c_v m}.$$

And now we can write out all of our terms together: putting together the $\delta(c_v)$ terms between all of our calculations will make things simplify further, and we end up with

$$g_1^{\overline{v}}(\tau) = \frac{2\pi i c_v}{N^2} + \delta(c_v)\zeta^{d_v}(1) + \frac{C_1}{N}\sum_{k=0}^{\infty}\sigma_0^{\overline{v}}(k) q_N^k = \boxed{G_1^{\overline{v}}(\tau) - \frac{C_1}{N}\left(\frac{c_v}{N} - \frac{1}{2}\right)},$$

where

$$\sigma_0^{\overline{v}}(k) = \sum_{m|k, k/m \equiv c_v}^{\infty} \mu_N^{md_v}$$

and

$$G_1^{\overline{v}}(\tau) = \delta(c_v)\zeta^{d_v}(1) + \frac{C_1}{N}\left(\sum_{k=1}^{\infty}\sigma_0^{\overline{v}}(k) q_N^k\right).$$

So we've now defined g_1 and G_1, and because the Fourier coefficients for g_1 are bounded, this is actually a modular form with respect to $\Gamma(N)$.

From here, we can define the series with respect to Dirichlet characters as well: if $uv = N$, ψ is a primitive character mod u, and ϕ is a character mod v, we can define

$$G_1^{\psi,\phi}(\tau) = \sum_{c=0}^{u-1}\sum_{d=0}^{v-1}\sum_{e=0}^{u-1} \psi(c)\overline{\phi}(d) g_1^{\overline{(cv, d+ev)}}(\tau)$$

and we'll define

$$E_1^{\psi,\phi}(\tau) = \delta(\phi)L(0,\psi) + \delta(\psi)L(0,\phi) + 2\sum_{n=1}^{\infty}\sigma_0^{\psi,\phi}(n) q^n.$$

We're claiming that these objects are actually the same up to a constant:

$$G_1^{\psi,\phi} = \frac{C_1 g(\overline{\phi})}{v} E_1^{\psi,\phi},$$

where $g(\overline{\phi})$ is the Gauss sum. In fact, we will establish a result similar to that of previous classes:

Theorem 296

Let $A_{n,1}$ be the set of triples $(\{\phi,\psi\}, t)$ (**unordered** pairs here) such that $tuv|N, \psi\phi(-1) = -1$, and ψ, ϕ are both primitive (mod u and v, respectively). Define $E_1^{\psi,\phi,t}(\tau) = E_1^{\psi,\phi}(t\tau)$. Then $\{E_1^{\{\psi,\phi\},t} : (\{\psi,\phi\}, t) \in A_{N,1}\}$ form a basis for the Eisenstein space of $\Gamma_1(N)$, and for any character χ,

$$\{E_1^{\psi,\phi,t} : (\{\phi,\psi,t) \in A_{N,1}, (\phi\psi) = \chi\}$$

forms a basis for the χ-eigenspace.

Proof. Everything except the **constant terms** actually work out exactly like our earlier proofs, because G_1 and E_1 are the same as the $k \geq 3$ definitions except for the constants. But finding this constant is not very easy. We'll break down the constant term for $\sum g_1$ into the sums for G_1 (which behaves like G_k for larger k) and $(g_1 - G_1)$. Summing over c, d, e means we can simplify the $(g_1 - G_1)$ term to

$$-\frac{C_1}{N}\sum_{c=0}^{u-1}\psi(c)\left(\frac{c_v}{N}-\frac{1}{2}\right)\sum_{d=0}^{v-1}\overline{\phi}(d)\sum_{e=0}^{u-1}1.$$

But now the inner two sums are zero unless we have ϕ being the trivial character $\mathbf{1}$, in which case $v = 1$ and this all simplifies to

$$= -C_1\delta(\phi)\sum_{c=0}^{N-1}\psi(c)\left(\frac{c_v}{N}-\frac{1}{2}\right) = -\frac{C_1 g(\overline{\phi})}{v}\delta(\phi)B_{1,\psi} = \frac{C_1 g(\overline{\phi})}{v}\delta(\phi)L(0,\psi)$$

by using properties from Michael's lecture.

The idea here is that when $\phi = 1$, we can multiply by things like $g(\overline{\phi})$ and v which are actually just 1, and then we can simplify to a form that we want.

On the other hand, the triple sum constant term has already been computed: it's going to be

$$\psi(0)\sum_{d=0}^{v-1}\overline{\phi}(d)\zeta^{\overline{d}}(1).$$

This is similarly zero except when ψ is the 1 character, and we'll write this as a limit as s approaches 1 of

$$\sum_{d=0}^{v-1}\overline{\phi}(d)\zeta^{\overline{d}}(s) = \sum_{d}\sum_{n\equiv d \bmod v}{}^{'}\,\overline{\phi}(n)n^{-s} = (1-(-1)^{-s})L(s,\overline{\phi}) = 2\delta(\psi)L(1,\overline{\phi}).$$

Note that we have a functional equation for the L-function, so $L(1,\overline{\phi})$ can be written in terms of $L(0,\phi)$ (in general, we can write $L(k,\phi)$ in terms of $L(1-k,\phi)$). So we end up with a total constant term

$$\frac{C_1 g(\overline{\phi}}{v}(\delta(\psi)L(0,\psi)+\delta(\psi)L(0,\phi)).$$

And this is exactly the constant scaling that we wanted when we defined $E_1^{\psi,\phi}$. □

23 May 7, 2020

Diamond and Shurman 4.9 – Dhruv Rohatgi

We'll be discussing the Fourier and Mellin transforms, and there won't really be very much number theory in this lecture. The idea is that we've claimed a lot of holomorphic and meromorphic functions can be analytically continued – we proved this using a functional equation for the Gamma function, but we didn't really do it rigorously for the zeta function. So we'll be discussing the ξ function, defined as

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

and we'll show that ξ has a meromorphic continuation to all of $\mathbb{C}$, which will show that ζ does as well.

Recall that the (ℓ-dimension) **theta function**

$$\theta(\tau,\ell) = \sum_{n\in\mathbb{Z}^\ell} e^{\pi i|n^2|\tau}$$

is holomorphic, because we're bounded away from the real axis for any compact subset of $\mathbb{C}$. We'll use a Poisson summation to prove a functional equation, which will allow us to construct and understand ξ.

Theorem 297

For all $t > 0$ and integer $\ell > 0$, we have some symmetry along the imaginary axis:

$$\theta(it, \ell) = \sum_{n \in \mathbb{Z}^\ell} e^{-\pi |n^2| t} = t^{-\ell/2} \sum_{n \in \mathbb{Z}^\ell} e^{-\pi |n^2|/t} = t^{-\ell/2} \theta(i/t, \ell).$$

We'll get a functional equation

$$\theta(-1/\tau, \ell) = (-i\tau)^{\ell/2} \theta(\tau, \ell)$$

as a corollary of this result.

In order to prove this functional equation, we'll need to introduce the Fourier transform:

Definition 298

The **Fourier transform** of an absolutely integrable $f : \mathbb{R}^\ell \to \mathbb{C}$ is defined to be

$$\hat{f}(y) = \int_{\mathbb{R}^\ell} f(x) e^{-2\pi i \langle x, y \rangle} dy.$$

Proposition 299

For "sufficiently nice" $f : \mathbb{R}^\ell \to \mathbb{C}$, we have

$$\sum_{n \in \mathbb{Z}^\ell} f(n) = \sum_{m \in \mathbb{Z}^\ell} \hat{f}(m).$$

Lemma 300

The Fourier transform of a Gaussian is another Gaussian with a scaling factor: if $f(x) = e^{-\pi t |x|^2}$ is a function from $\mathbb{R}^\ell \to \mathbb{R}$, then

$$\hat{f}(y) = t^{-\ell/2} e^{-\pi |y|^2/t}.$$

Proof. This is a "contour shifting" proof:

$$\hat{f}(y) = \int_{\mathbb{R}^\ell} e^{-\pi t |x|^2} e^{-2\pi i \langle x, y \rangle dx} = e^{-\pi |y|^2/t} \int_{\mathbb{R}^\ell} e^{-\pi t |x + iy/t|^2} dx.$$

But integrating over a shifted axis is the same as the regular integral of a Gaussian (by Cauchy's theorem), and we can evaluate it to get the desired result. □

Applying this lemma and the Poisson summation formula gives us exactly what we want to prove the above theorem.

Definition 301

Let $f : \mathbb{R}^+ \to \mathbb{C}$ be a function. Then the **Mellin transform** of f is

$$g(s) = \int_0^\infty f(t) t^{s-1} dt.$$

This is nice because it has an inverse (and other nice properties), but we mostly care about it for motivation. Notably, if we pick $f(t) = \frac{1}{2}(\theta(it) - 1)$, then

$$g(s) = \frac{1}{2} \int_0^\infty (\theta(it) - 1) t^{s-1} dt.$$

Lemma 302

When $\mathrm{Re}(s) > 1/2$, $g(s)$ is well-defined, and

$$g(s) = \pi^{-s}\Gamma(s)\zeta(2s) = \xi(2s).$$

Proof. We check that we don't have "problem points" at $t \to 0$ (since $\theta(it)$ might explode) or $t \to \infty$ (since t^{s-1} might explode). Break up the integral into a part from 0 to 1 and a part from 1 to ∞. The latter decays exponentially because $|\theta(it) - 1| \le 2\sum_{n=1}^{\infty} e^{-\pi n^2 t}$, and to understand the former, use the functional equation to find that

$$\theta(it) = t^{-1/2}\theta(i/t),$$

so $\theta(i/t)$ goes to 1 as $t \to 0$, meaning $\theta(it)$ is asymptotically $t^{-1/2}$. This means that when we multiply it by t^{s-1} for any $\mathrm{Re}(s) > 1/2$, this integral does indeed converge.

To show that this is related to the ξ function, we can write out the definition of the theta function:

$$g(s) = \frac{1}{2}\int_0^\infty (\theta(it) - 1)t^{s-1}dt = \int_0^\infty \sum_{n=1}^{\infty} e^{-\pi n^2 t}t^{s-1}dt,$$

where we've summed over positive integers instead of all integers and removed the factor of 2. We can then use the dominated convergence theorem to swap the sum and integral to get

$$= \sum_{n=1}^{\infty}\int_0^\infty e^{-\pi n^2 t}t^{s-1}dt.$$

Substituting $u = \pi n^2 t$, the integral becomes independent of n, and we end up with

$$= \sum_{n=1}^{\infty}(\pi n^2)^{-s}\int_0^\infty e^{-u}u^{s-1}du,$$

which miraculously simplifies to $\pi^{-s}\zeta(2s)\Gamma(s)$, as desired. □

So we've shown that we have another expression for ξ on the same domain it's defined, and now we can show that g extends meromorphically to $\mathbb{C}$ – this is easier than working with ξ directly. We again split into its two pieces:

$$g(s) = \int_0^1 (\theta(it) - 1)t^{s-1}dt + \int_1^\infty (\theta(it) - 1)t^{s-1}dt.$$

The second integral is already holomorphic, so we don't need to worry about it. The first integral is essentially going to look asymptotically like

$$\approx \int_0^1 (t^{-1/2} - 1)t^{s-1}dt = -\frac{1}{1/2 - s} - \frac{1}{s},$$

and notice that the function on the right side is indeed meromorphic everywhere on $\mathbb{C}$. So we can analyze in terms of an "error term:" whenever $\mathrm{Re}(s) > \frac{1}{2}$, we can first integrate the polynomial term

$$\int_0^1 (\theta(it) - 1)t^{s-1}dt = \int_0^1 \theta(it)t^{s-1}dt - \frac{1}{s},$$

and then do a u-substitution $t = \frac{1}{u}$ to get

$$g(s) = \int_1^\infty (\theta(it) - 1)(t^{-s-1/2} + t^{s-1})dt - \frac{1}{s} - \frac{1}{1/2 - s}.$$

But this extra term is meromorphic, and that's the same thing as saying that g (and therefore ξ meromorphically continue). And notice that there's a symmetry in this expression: $g(s) = g\left(\frac{1}{2} - s\right)$, and since $g(s) = \theta(2s)$, this means that $\xi(s) = \xi(1-s)$. So we've shown that ξ extends meromorphically to all of $\mathbb{C}$, with only simple poles at 0 and 1, with a simple functional equation. This allows us to understand the poles of ζ pretty clearly as well (it's just at $z = 1$).

Diamond and Shurman 4.10 – Swapnil Garg

In this lecture, we'll meromorphically extend the Eisenstein series using the Mellin transform we just introduced. The idea here is that in Serre, we discussed the Eisenstein series for $SL_2(\mathbb{Z})$ for all even weight $k \ge 4$, and similarly we do this for $k \ge 3$ for congruence subgroups (where the weight doesn't necessarily have to be even). Eisenstein series also exist for $k = 1, 2$, but they're harder to deal with because they don't converge. There's a lot more bashing that we have to do, but we'll look at these series directly by adding a complex parameter s. We will then analytically continue our series to $s = 0$, which is theoretically what our series should look like.

Definition 303

Let $\overline{v}$ be a vector in $(\mathbb{Z}/N\mathbb{Z})^2$ of order N, let k be a positive integer, and ε_N be defined as usual ($\frac{1}{2}$ for $N = 1, 2$ and 1 otherwise). Let v be a lift of $\overline{v}$ to $\mathbb{Z}^2$. Then for any $\tau = (x + iy) \in \mathbb{H}$ and $s \in \mathbb{C}$, define the **augmented series** via

$$E_k^{\overline{v}}(\tau, s) = \varepsilon_N \sum_{\substack{(c,d) \equiv v \bmod N \\ \gcd(c,d)=1}} \frac{\text{Im}(\tau)^s}{(c\tau + d)^k |c\tau + d|^{2s}}.$$

Note that whenever $\text{Re}(k + 2s) > 2$, this series converges absolutely, and because it converges uniformly on compact sets, this makes the series analytic on that particular half-plane.

We're going to rewrite this series using weight-k operators: we'll **generalize them to two parameters** by letting

$$(f[\gamma]_k)(\tau, s) = j(\gamma, \tau)^{-k} f(\gamma(\tau), s),$$

where $j(\gamma, \tau)$ is still $c\tau + d$. We showed earlier on that some functions f satisfy $f[\gamma]_k = f$, meaning that we are weakly modular of weight k. So if we let δ be an element of $SL_2(\mathbb{Z})$ with bottom row (c_v, d_v), and we let P_+ be the positive part of the parabolic subgroup of $SL_2(\mathbb{Z})$ (meaning we consider the elements $\begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix}$, where n is a positive integer), we have the following result:

Proposition 304

The Eisenstein series can be rewritten as

$$E_k^{\overline{v}}(\tau, s) = \varepsilon_N \sum_{\gamma \in (P_+ \cap \Gamma(N)) \backslash \Gamma(N)\delta} \text{Im}(\tau)^s [\gamma]_k$$

(where we sum over the orbits of $P_+ \cap \Gamma(N)$ inside $\Gamma(N)\delta$).

Proof. Compare this expression to the original definition of $E_k^{\overline{v}}(\tau, s)$. First, we can check that for any $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ in $SL_2(\mathbb{Z})$, we have

$$\text{Im}(\gamma(\tau)) = \frac{\text{Im}(\tau)}{|c\tau + d|^2},$$

so that means the $\text{Im}(\tau)^s$ and $|c\tau + d|^{2s}$ cancel out the relevant factors. Looking at the stabilizer of $\Gamma(N)\delta$, two matrices in $\Gamma(N)\delta$ will have the same bottom row if and only if they are in the same orbit of $P_+ \cap \Gamma(N)$, because $\Gamma(N)\delta$ only contains matrices with bottom row $(c, d) \equiv v$. This means that we know what the weight-k operator does:

$$\text{Im}(\tau)^s[\gamma]_k = (c\tau + d)^{-k}\,\text{Im}(\gamma(\tau))^s = \frac{\text{Im}(\tau)^s}{(c\tau + d)^k|c\tau + d|^{2s}},$$

and plugging this back in and looking at what we sum over (each orbit corresponds to a specific vector (c, d)) yields the result. □

This yields the following formula, which should look familiar:

Lemma 305

We have

$$(E_k^{\overline{v}}[\gamma]_k)(\tau, s) = E_k^{\overline{v\gamma}}(\tau, s).$$

Proof. This is the same as the proof in Section 4.2: we just sum over the orbits in a different order and gain an extra factor. □

In particular, for any $\gamma \in \Gamma(N)$, it is true that $(E_k^{\overline{v}}[\gamma]_k)(\tau, s) = E_k^{\overline{v}}(\tau, s)$ (because $\overline{v\gamma} = \overline{v}$). TO proceed, we can look at the non-normalized Eisenstein series

$$G_k^{\overline{v}}(\tau, s) = \sum_{(c,d)\equiv v \mod N} \frac{\text{Im}(\tau)^s}{(c\tau + d)^k|c\tau + d|^{2s}}.$$

Remember that G and E are basically the same – there are some linear combinations – so if we can meromorphically continue one, we can meromorphically continue the other. So we'll **only look at the case** $k = 0$ and $N = 1$ today, which is very nice, and we'll see the general case next time.

So now we can look at the **modified theta function**

$$\vartheta(\gamma) = \sum_{n\in\mathbb{Z}^2} e^{-\pi|n\gamma|^2}$$

where $\gamma \in GL_2(\mathbb{R})$.

Lemma 306

For a function $f \in L^1(\mathbb{R}^2)$, a matrix $\gamma \in SL_2(\mathbb{R})$, and $r > 0$, the Fourier transform of a general function (for $x \in \mathbb{R}^2$) $\phi(x) = f(x\gamma r)$ is $\hat{\theta}(x) = r^{-2}\hat{f}(x\gamma^{-T}/r)$, where γ^{-T} denotes the negative transpose for γ.

We're skipping over this proof for now – it's just a fact about Fourier transforms. So now if we use Poisson summation and use the function $f(x) = e^{-\pi|x|^2}$, we find that

$$r\sum_{n\in\mathbb{Z}^2} f(n\gamma r) = r^{-1}\sum_{n\in\mathbb{Z}^2} f(n\gamma^{-T}/r).$$

for any $r > 0$ and $\gamma \in SL_2(\mathbb{R})$. So now if we let $S = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, then $S\gamma^{-T} = \gamma S$, and $|x|$ and $|xS|$ are the same. Since f is a Gaussian function, this means $f(x) = f(xS)$, and S is just a 90 degree rotation of our lattice, so we sum over the same thing. Thus we get the following result:

Corollary 307

We have the **transformation law** by using the fact that $\vartheta(\gamma) = \sum_{n\in\mathbb{Z}^2} e^{-\pi|n\gamma|^2}$.

$$r\vartheta(\gamma r) = \frac{1}{r}\vartheta\left(\frac{\gamma}{r}\right).$$

Now we look at something similar to in Dhruv's lecture: note that the Mellin transform of $\vartheta(\gamma t^{1/2}) - 1$,

$$g(s,\gamma) = \int_0^\infty (\vartheta(\gamma t^{1/2} - 1))t^s \frac{dt}{t} = \int_t^\infty \sum_{n\in\mathbb{Z}^2}' e^{-\pi|n\gamma|^2 t} t^s \frac{dt}{t}$$

(note that letting $\gamma = iI$ makes this the same as the previous lecture), where subtracting the 1 in the first expression removes the point at the origin (which is why we have a $\sum'$). The transformation law tells us that $\vartheta(\gamma t^{1/2})$ must be asymptotically proportional to $\frac{1}{t}$ near $t = 0$, because our function goes to 1 at infinity. That means that our integral for $g(s,\gamma)$ converges at $t = 0$ as long as $\text{Re}(s) > 1$. But this function also converges at all s for $t \to \infty$ (again, by exponential suppression). So we can move the sum outside the integral because of convergence, and changing variables with t replacing $\pi|n\gamma|^2 t$ yields our gamma function again:

$$g(s,\gamma) = \sum_{n\in\mathbb{Z}^2}' (\pi|n\gamma|^2)^{-s} \int_0^\infty e^{-t} t^s \frac{dt}{t} = \pi^{-s}\Gamma(s) \sum_{n\in\mathbb{Z}^2}' |n\gamma|^{-2s}.$$

So now we can connect this to the function $G_0(\tau, s)$: for any $\tau = x + iy$ in the upper half-plane, define

$$\gamma_\tau = \frac{1}{\sqrt{y}} \begin{bmatrix} y & x \\ 0 & 1 \end{bmatrix},$$

which sends i to τ. Since $|(c,d)\gamma_\tau| = |c\tau + d|^2/y$ (just by computation), we find that

$$g(s,\gamma_\tau) = \pi^{-s}\Gamma(s) \sum_{(c,d)\in\mathbb{Z}^2}' \frac{y^s}{|c\tau + d|^{2s}},$$

which is exactly the expression $\pi^{-s}\Gamma(s)G_0(\tau,s)$ for the augmented Eisenstein series when $k = 0$ and $N = 1$. A bit of algebra analogous to the last section (a change of variables), plus the transformation law, yields

$$\int_0^1 (\vartheta(\gamma t^{1/2}) - 1)t^s \frac{dt}{t} = \int_1^\infty (\vartheta(\gamma t^{1/2}) - 1)t^{1-s} \frac{dt}{t} - \frac{1}{s} - \frac{1}{1-s}.$$

But notice that this integrand converges quickly – there are no problem points – which shows that $g(s,\gamma)$ only has simple poles at $s = 0$ and 1, and we have $g(s,\gamma) = g(1-s,\gamma)$. So

$$\pi^{-s}\Gamma(s)G_0(\tau,s) = \pi^{s-1}\Gamma(1-s)G_0(\tau,1-s),$$

and we've defined our weight-0 Eisenstein series (in a simple case) with a meromorphic continuation. This function is $SL_2(\mathbb{Z})$-invariant and of weight 0, so it is weakly modular.

24 May 12, 2020

Diamond and Shurman 4.10 continued – Shreyas Balaji

We'll start with some definitions from last lecture: we have

$$\vartheta(\gamma) = \sum_{n\in\mathbb{Z}^2} e^{-\pi|n\gamma|^2}, \gamma \in GL_2(\mathbb{R}),$$

$$\gamma_\tau = \frac{1}{\sqrt{y}}\begin{bmatrix} y & x \\ 0 & 1 \end{bmatrix} \in SL_2(\mathbb{R}), \quad \tau = x + iy,$$

$$G_k^{\overline{v}}(\tau, s) = \sum_{c,d\equiv v \bmod N} \frac{y^s}{(c\tau+d)^k|c\tau+d|^{2s}}.$$

Last time, we worked with these objects by considering the function $\vartheta(\gamma t^{1/2}) - 1$, and we considered the Mellin transform $g(s,\gamma)$, which was both equal to $\pi^{-s}\Gamma(s)G_0(\tau, s)$ and an meromorphic function in s

$$g(s,\gamma) = \int_1^\infty (\vartheta(\gamma t^{1/2}) - 1)(t^s + t^{1-s})\frac{dt}{t} - \frac{1}{s} - \frac{1}{1-s}$$

which has a meromorphic extension to the full complex plane. Our goal today is extend this logic to higher weights and levels k, N.

We'll introduce some notation: let G be $(\mathbb{Z}/N\mathbb{Z})^2$ (we'll be treating these group elements as row vectors), let $\mu_N = e^{2\pi i/N}$, and we'll be considering arbitrary functions $a : G \to \mathbb{C}$ instead of vectors (they don't need to be homomorphisms). We'll let $\langle \cdot, \cdot \rangle$ denote the standard inner product, and we'll also define $S = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. Note that $S^T = -S$, which is nice for inner products which we'll be working with throughout this lecture.

Definition 308

The **Fourier transform** of a function $a : G \to \mathbb{C}$ is given by

$$\hat{a}(\overline{v}) = \frac{1}{N}\sum_{\overline{w}\in G} a(\overline{w})\mu_N^{-\langle w, vS\rangle}.$$

Proposition 309

We can invert the Fourier transform to get

$$a(\overline{u}) = \frac{1}{N}\sum_{\overline{v}\in G} \hat{a}(\overline{v})\mu_N^{\langle u, vS\rangle}$$

Proof. There is some basis vector e_j such that $\mu_N^{\langle u, e_j S\rangle} \neq 1$ as long as $\overline{u}$ is nonzero, and then we can "sum over the group" in two ways by the transformation $v \to v + e_j$, showing that the sum $\sum_{v\in G}\mu_N^{\langle u, vS\rangle}$ will be 0 for all $\overline{u} \neq 0$. (And when $\overline{u} = 0$, that sum is N^2.) So we can plug in the definition for $\hat{a}(\overline{v})$ and swap the order of summation, which will yield the result. □

Definition 310

The **harmonic polynomial** for a positive integer k, denoted h_k, is defined as

$$h_k(c, d) = (-i)^k(c + id)^k$$

for $(c, d) \in \mathbb{R}^2$.

Definition 311

Define the **theta function** $\theta_k^{\overline{v}} : GL_2(\mathbb{R}) \to \mathbb{C}$ to be

$$\vartheta_k^{\overline{v}}(\gamma) = \sum_{n\in\mathbb{Z}^2} h_k((v/N + n)\gamma)e^{-\pi|(v/N+n)\gamma|^2}.$$

For notation's sake, let $f(x) = e^{-\pi|x|^2}$ be the Gaussian function and $f_k(x) = h_k(x)e^{-\pi|x|^2}$ be the Schwartz function. There's lots of useful facts we'll need before doing some algebraic manipulation:

- The Fourier transform of the Schwartz function f_k is $(-i)^k f_k$.
- Define $\phi_k(x) = f_k(x\gamma r)$ for some $r > 0$: then the Fourier transform is $(-i)^k r^{-2} f_k(x\gamma^{-T}r^{-1})$.
- Expanding out the definition of h_k and defining $z(x) = c + id$ for $x = (c, d)$, we can write

$$h_k(x) = (z(xS)^k.$$

- This means that $h_k(xS) = (-i)^k h_k(x)$, and $f_k(xS) = (-i)^k h_k(x)$.
- Finally, $S\gamma^{-T} = \gamma S$ and $S^T = -S$.

So now we have a lot of algebra to get through: we expand out $r\vartheta_k^{\overline{v}}(\gamma r)$ from first definitions and write it in terms of the Schwarz function and then ϕ_k (by definition) to get

$$= r \sum_{n\in Z\mathbb{Z}^2} f_k((v/N + n)\gamma r) = r \sum_{n\in\mathbb{Z}^2} \phi_k(v/N + n).$$

Recall the Poisson summation formula: applying it to our function (where we use the Fourier transform) yields

$$= r \sum_{n\in\mathbb{Z}^2} \hat{\phi}_k(n)e^{2\pi i\langle n, v/N\rangle} = (-i)^k r^{-1} \sum_{n\in\mathbb{Z}^2} f_k(n\gamma^{-T}r^{-1})e^{2\pi i\langle n, v/N\rangle}.$$

Substituting $n \to nS$ and swapping around a few terms yields

$$= (-i)^k r^{-1} \sum_{n\in\mathbb{Z}^2} (-i)^k f_k(n\gamma r^{-1})\mu_N^{-\langle n, vS\rangle},$$

and from here, we break up our sum by group elements in G:

$$= (-1)^k r^{-1} \sum_{\overline{w}\in G} \sum_{n\in\mathbb{Z}^2, n\equiv w \mod N} f_k(n\gamma r^{-1})\mu_N^{-\langle w, vS\rangle},$$

and now this is exactly the sum that we want in the definition of our theta function: some more algebra yields

$$= (-1)^k r^{-1} \sum_{\overline{w}\in G} \vartheta_k^{\overline{w}}(\gamma N r^{-1})\mu_N^{-\langle w, vS\rangle}.$$

If we now look at $\vartheta^{\overline{v}}_k$ as a function of $\overline{v}$, then our last sum looks like N times the Fourier transform. Adding that factor back in yields

$$r\theta^{\overline{v}}_k(\gamma r) = (-1)^k N r^{-1}\hat{\vartheta}^{\overline{v}}_k(\gamma N r^{-1}).$$

And since r is arbitrary, we send $r \to N^{1/2}r$, and we get the final identity

$$\boxed{r\vartheta^{\overline{v}}_k(\gamma N^{1/2}r) = (-1)^k r^{-1}\hat{\vartheta}^{\overline{v}}_k(\gamma N^{1/2}r^{-1})}.$$

Proposition 312

Let $a : G \to \mathbb{C}$ be a function. Then the **sum-theta** function

$$\Theta_k(\gamma) = \sum_{\overline{v}\in G}(a\overline{v} + (-1)^k\hat{a}(-\overline{v}))\vartheta^{[}_k\overline{v}(\gamma N^{1/2})$$

satisfies $r\Theta^a_k(\gamma r) = r^{-1}\Theta^a_k(\gamma r^{-1})$.

To show this, we'll need a lemma that relates the Fourier transforms with swapping $v \to -v$:

Lemma 313

For any function $a : G \to \mathbb{C}$, we have

$$\sum_{\overline{v}\in G} a(\overline{v})\vartheta^{\overline{v}}_k(\gamma) = \sum_{\overline{v}\in G}\hat{a}(-\overline{v})\hat{\vartheta}^{\overline{v}}_k(\gamma),$$

$$\sum_{\overline{v}\in G}\hat{a}(-\overline{v})\vartheta^{\overline{v}}_k(\gamma) = \sum_{\overline{v}\in G} a(\overline{v})\hat{\vartheta}^{\overline{v}}_k(\gamma).$$

Proof of lemma. To do this, we can prove the general statement

$$\sum_{x\in G} a(x)b(x) = \sum_{y\in G}\hat{a}(-y)\hat{b}(y)$$

by writing out the Fourier transforms and swapping the order of summation. Then substituting in $b = \vartheta^{\overline{v}}_k$ or $\hat{\vartheta}^{\overline{v}}_k$ yields the desired result. □

Proof of Proposition 312. Write out the definition of sum-theta function, apply the boxed identity above, and then use the lemma to move factors of -1 between the a functions. □

This is a generalization of what was shown in the last lecture. Note that $\Theta^a_k(\gamma r)$ rapidly converges to 0 as $r \to \infty$, so the following Mellin transform of Θ^a_k, denoted

$$g^a_k(s,\gamma) = \int_0^\infty \Theta^a_k(\Gamma t^{1/2})t^s\frac{dt}{t},$$

converges as $t \to 0$ for all s because of the proposition which essentially allows us to "swap" $t^{1/2}$ with $t^{-1/2}$.

More specifically, the integral converges as $t \to 0$, as well as $t \to \infty$. So we can expand the sum, swapping the sum and the integral, expanding out the expression for the theta function, and then factor out all the terms that we can that don't depend on t, using the trick where we sum over all $n \equiv v \bmod N$ instead of all $n \in \mathbb{Z}^2$. The final result we find is that

$$g_k^a(s,\gamma) = \pi^{-k/2-s} N^s \Gamma(k/2+s) \sum_{\overline{v}\in G} (a(\overline{v}) + (-1)^k \hat{a}(-\overline{v})) \sideset{}{'}\sum_{n\equiv v \bmod N} h_k(n\gamma)|n\gamma|^{-k-2s}.$$

Now defining $\gamma = \gamma_\tau$ yields $h_k(n\gamma_\tau) = (\overline{c\tau+d})^k/y^{k/2}$, and we also have that $|n\gamma|^{-k-2s} = y^{k/2+s}/|c\tau+d|^{k+2s}$, so we can substitute this in and simplify to find that the Mellin transform of the sum-theta function (specifying to γ_τ) is

$$g_k^a(s,\gamma_\tau) = \pi^{-k/2-s}\Gamma(k/2+s)N^s y^{k/2} G_k^a(\tau, s-k/2),$$

where G_k^a is the sum of Eisenstein series given by

$$G_k^a(\tau,s) = \sum_{\overline{v}\in G} (a(\overline{v}) + (-1)^k \hat{a}(-\overline{v})) G_k^{\overline{v}}(\gamma N^{1/2}).$$

And now we do the same thing with analytic continuation: applying our above proposition on the Mellin transform integral lets us relate values of Θ_k^a between the regions $[0,1]$ and $[1,\infty]$, and the transformation yields

$$\int_0^1 \Theta_k^a(\gamma t^{1/2}) t^s \frac{dt}{t} = \int_1^\infty \Theta_k^a(\gamma t^{1/2})(t^s + t^{1-s})\frac{dt}{t},$$

which has symmetry under $s \to (1-s)$. And this last integral is entire in s, so combining everything gives us the final theorem of the section:

Theorem 314

For any positive integer N, let $G = (\mathbb{Z}/N\mathbb{Z})^2$. Construct $G_k^a(\tau,s)$ as before: then for any integer k and any point $\tau = x+iy \in \mathbb{H}$,

$$(\pi/N)^{-s}\Gamma(|k|/2+s)G_k^a(\tau, s-k/2)$$

has a meromorphic continuation in s to all of $\mathbb{C}$ which is entire when $k \neq 0$ and has simple poles at $0, 1$ for $k = 0$.

(Specifically, we've proved the $k > 0$ case, and the others follow by a similar argument.)

Diamond and Shurman 4.11 – Anton Trygub

We'll make this presentation as short as possible to leave some time for celebration at the end of the class. Basically, we'll introduce a new theta function and show a few important properties.

Lemma 315

Let d be a cubefree positive integer. The equation $x^3 = d \bmod p$ has 3 solutions if $p \equiv 1 \bmod 3$ and d is a nonzero cube mod p, 0 solutions if $p \equiv 1 \bmod 3$ and d is not a cube mod p, and 1 if $p \equiv 2 \bmod 3$ or $p|3d$.

Proof. The result is clear for $p = 3$ or $p = d$ (because $0^3, 1^3, 2^3$ are all different mod 3, and $d = 0$ is the only solution in the latter case). otherwise, let g be a primitive root mod p: letting $d = p^k$, the equation reduces to showing the number of solutions to $3x = k \bmod (p-1)$, and the result follows. □

We'll denote $e(z) = e^{2\pi i z}$ and $\text{tr}(x) = x + x^*$ throughout the next few results. We'll also use the notation $A = \mathbb{Z}[\mu_3], \alpha = i\sqrt{3}, B = \frac{1}{\alpha}A$ (note that $A \subset B \subset \frac{1}{3}A$ as lattices). Some useful facts to remember are that if we write $x = x_1 + x_2\mu$ for integers x_1, x_2,

$$|x|^2 = x_1^2 - x_1x_2 + x_2^2, \quad |x+y|^2 = |x|^2 + |y|^2 + \text{tr}(xy^*).$$

Definition 316

Let N be a positive integer, and let $\overline{u} \in \frac{1}{3}A/NA$. Then we have the theta function

$$\theta^{\overline{u}}(\tau, N) = \sum_{n \in A} e\left(N \left| \frac{u}{N} + n \right|^2 \tau \right)$$

for $\tau \in \mathbb{H}$.

Our first result tells us about this function when we move along the real axis:

Lemma 317

We have

$$\theta^{\overline{u}}(\tau + 1, N) = e\left(\frac{|u|^2}{N} \right) \theta^{\overline{u}}(\tau, N)$$

for any positive integer N and $\overline{u} \in B/NA$.

Proof. We have that

$$N \left| \frac{u}{N} + n \right|^2 \equiv \frac{|u|^2}{N} \text{ mod } \mathbb{Z}$$

by expanding out $u = u_1 + u_2\mu$ and $n = n_1 + n_2\mu$. Since $e(a+b) = e(a)e(b)$ by definition, replacing τ with $\tau + 1$ in the definition will give us an extra $e\left(N \left| \frac{u}{N} + n \right|^2 \right)$, and this simplifies using our first observation because e is 1-periodic in the real direction. □

The next result is a "scaling:"

Lemma 318

We have

$$\theta^{\overline{u}}(\tau, N) = \sum_{\substack{\overline{v} \in B/(dNA) \\ \overline{v} \equiv \overline{n} \text{ mod } NA}} \theta^{\overline{v}}(dr, dN)$$

for any positive integer N, $\overline{u} \in B/NA$, and positive integer d.

Proof. Let $n = r + dm$, and substitute into the definition:

$$\theta^{\overline{u}}(\tau, N) = \sum_{n \in A} e\left(N \left| \frac{u}{N} + n \right|^2 \tau \right) = \sum_{\substack{m \in A \\ \overline{r} \in A/dA}} e\left(dN \left| \frac{u + rN}{dN} + m \right|^2 d\tau \right)$$

where we've introduced a factor of $\frac{1}{d}$ inside the absolute value square but canceled this out with the two factors of d outside. But now the sum over m can be rewritten as a theta function as well, which yields what we want. □

Finally, we describe how θ transforms under the "action" of S:

Lemma 319

We have

$$\theta^{\overline{u}}\left(-\frac{1}{\tau}, N \right) = \frac{-i\tau}{N\sqrt{3}} \sum_{\overline{v} \in B/NA} e\left(-\frac{\operatorname{tr}(uv^*)}{N} \right) \theta^{\overline{v}}(\tau, N)$$

for any positive integer N and $\overline{u} \in B/NA$.

This result takes a bit more effort to prove, so we'll leave it as an exercise.