6.1 p-adic Modular Forms

In the first half of this Arizona Winter Semester, you learned about *modular forms*. Here we focus on the modular forms for the full modular group $\Gamma = SL_2(\mathbb{Z})$. A modular form can be studied via its Fourier series, which can be viewed as a formal power series in q. Modular forms are graded by weight: given a coefficient ring $A \subset \mathbb{C}$ and weight k, we let $M_k(A)$ denote the space of modular forms of weight k whose Fourier series has coefficients in A. We can also consider the space of modular forms of all weights, $M(A) = \oplus_k M_k(A)$.

Now we can define p-adic modular forms by taking p-adic limits of modular forms in $M(\mathbb{Q})$:

Definition 6.1

A *p-adic modular form* is a formal power series $f \in \mathbb{Q}_p[[q]]$ such that there exists a sequence $(f_i) : f_i \in M(\mathbb{Q})$ whose coefficients converge uniformly to f.

Note that we have not said anything yet about what the weight of a p-adic modular form is, though we will get to that later.

For now, to get a first look at what p-adic modular forms might look like, we consider certain limits of the Eisenstein series. We use the following normalizations which have rational coefficients:

$$\mathbb{G}_k = -\frac{B_k}{2k} + \sum_{n\geq 1} \sigma_{k-1}(n)q^n$$

$$E_k = 1 - \frac{2k}{B_k}\sum_{n\geq 1} \sigma_{k-1}(n)q^n$$

Here B_k are the Bernoulli numbers and $\sigma_k(n) = \sum_{d|n} d^k$.

We can take a limit of the $\mathbb{G}_k$ as follows. We introduce a version of σ_k with the p-part removed:

$$\sigma_k^*(n) := \sum_{d|n, p\nmid d} d^k$$

Let $m \geq 1$ be a positive natural number. Since the values of d appearing in the sum are units in $\mathbb{Z}/p^m\mathbb{Z}$, and the group of units has order $p^{m-1}(p-1)$, it follows that

$$\sigma_k^*(n) \equiv \sigma_{k'}^*(n) \pmod{p^m} \quad \text{whenever} \quad k \equiv k' \pmod{p^{m-1}(p-1)}$$

So if (k_i) is a sequence of weights which is eventually stable in $\mathbb{Z}/p^{m-1}(p-1)\mathbb{Z}$, then $\sigma_{k_i}(n)$ is a uniformly Cauchy set of sequences in $\mathbb{Z}_p$ (here we use the fact that the p-part of $\sigma_k(n)$ eventually vanishes mod p^m).

This guarantees that the nonconstant part of the $\mathbb{G}_{k_i}$ converge in $\mathbb{Q}_p[[q]$. The convergence of the constant term requires some work, but as we will see later, it converges as well.

Note that the limit of such a sequence (k_i) can be understood to live in

$$\mathfrak{X} := \underset{\leftarrow}{\lim} \mathbb{Z}/p^{m-1}(p-1) = \underset{\leftarrow}{\lim} \mathbb{Z}/p^{m-1} \times \mathbb{Z}/(p-1) = \mathbb{Z}_p \times \mathbb{Z}/(p-1)$$

So if $k = \lim_{i\to\infty}(k_i) \in \mathfrak{X}$, we can define

$$\sigma_k^*(n) := \lim_{i\to\infty} \sigma_{k_i}^*(n)$$

which is well-defined by the above considerations, and we have

$$\mathbb{G}_k^* = \lim_{i\to\infty} -\frac{B_{k_i}}{2k_i} + \sum_{n\geq 1} \sigma_k^*(n) q^n$$

a p-adic modular form! In fact, the constant term of this series can be used to define the p-adic ζ function as you will see in this week's problem set.

Now we take a closer look at the structure of p-adic modular forms. We introduce some definitions.

- We'll let $A_{(p)}$ denote the subring of $\mathbb{Q}$

  $$A_{(p)} := \mathbb{Q} \cap \mathbb{Z}_p = \left\{ \frac{a}{b} \in \mathbb{Q} : p \nmid b \right\}$$

  (note: this is usually denoted $\mathbb{Z}_{(p)}$).
- Let $M_k(A_{(p)})$ denote the modular forms of weight k with coefficients in $A_{(p)}$.
- Suppose $f = \sum_{n=0}^{\infty} a_i q^i \in M_k(A_{(p)})$; we denote $\overline{f} := \sum_{n=0}^{\infty} \overline{a_i} q^i$ where $\overline{a_i}$ is the reduction of a_i mod p.
- Let $M_k(\mathbb{F}_p) \subset \mathbb{F}_p[[q]]$ denote the image under reduction mod p of $M_k(A_{(p)})$, so

  $$M_k(\mathbb{F}_p) = \{\overline{f} : f \in M_k(A_{(p)})\}$$

Let's get a clearer picture of the elements of $M_k(\mathbb{F}_p)$. We can use the plentiful results we have about modular forms built up in the first half of AWS.

Firstly, we can use results on the Bernoulli numbers to determine the image of Eisenstein series mod p. The Clausen–Von Staudt theorem tells us that if $p-1$ divides k, then $v_p(k/B_k) \geq 1$. So all nonconstant terms of E_{p-1} vanish mod p, and so

$$E_{p-1} \equiv 1 \pmod{p}.$$

We note that if $f \in M_k(A_{(p)})$, then fE_{p-1} has weight $k+p-1$, and since $\overline{E_{p-1}} = 1$,

$$\overline{f} = \overline{fE_{p-1}} \in M_{k+p-1}$$

Thus we have a chain of inclusions

$$M_k(\mathbb{F}_p) \subset M_{k+p-1}(\mathbb{F}_p) \subset M_{k+2(p-1)} \subset \cdots$$

If $\alpha \in \mathbb{Z}/(p-1)\mathbb{Z}$, we let $M^\alpha(\mathbb{F}_p) := \bigcup_{k \equiv \alpha \pmod{p-1}} M_k(\mathbb{F}_p)$, a union of such a chain.

The following theorem shows that the relation $\overline{E_{p-1}} = 1$ tells us all relations, and that the $M^\alpha(\mathbb{F}_p)$ provide a grading of $M(\mathbb{F}_p)$:

Theorem 6.2 Swinnerton-Dyer

For $p \geq 5$, we have

$$M(\mathbb{F}_p) = \mathbb{F}_p[\overline{E}_4, \overline{E}_6]/(\overline{E}_{p-1} - 1)$$

$$M(\mathbb{F}_p) = \bigoplus_{\alpha \in \mathbb{Z}/(p-1)\mathbb{Z}} M^\alpha(\mathbb{F}_p)$$

Next, we refine our study, from coefficients in $\mathbb{F}_p$ to coefficients in $\mathbb{Z}/p^m\mathbb{Z}$:

Theorem 6.3 Weight congruences for congruent forms

Let $f \in M_k(A_{(p)})$ and $f' \in M_{k'}(A_{(p)})$. If

$$f \equiv f' \pmod{p^m}$$

then

$$k \equiv k' \pmod{p^{m-1}(p-1)} \text{ if } p \geq 3$$
$$k \equiv k' \pmod{2^{m-2}} \text{ if } p = 2$$

The case of this theorem when $p \geq 5$ and $m = 1$ follows from Theorem 5.2.

The consequence of this Theorem is that we can make sense of the weight of a p-adic modular form! Indeed, if $f \in \mathbb{Q}_p[[q]]$ is a p-adic modular form, it must be the uniform limit of a sequence of rational modular forms f_i. These f_i eventually agree mod p^m for each m and hence their weights k_i eventually agree mod $p^{m-1}(p-1)$ for each m, thus converging to a limit in $\mathfrak{X}$. Given $k \in \mathfrak{X}$, we let $M_k(\mathbb{Q}_p)$ denote the space of modular forms of weight k.

The following theorem assures us that it suffices to establish uniform convergence for the nonconstant coefficients, so in particular the p-adic limit of the Eisenstein series discussed above indeed exists.

Proposition 6.4

Let f_i be rational modular forms of weights k_i such that $\lim_{i\to\infty} k_i = k \in \mathfrak{X}$ and the *nonconstant* Fourier coefficients of the f_i converge uniformly in $\mathbb{Z}_p$. Then the constant coefficients converge in $\mathbb{Z}_p$ as well, so that the f_i converge to a p-adic modular form $f \in M_k(\mathbb{Q}_p)$.

Moreover, if $f = \sum_{n=0}^{\infty} a_n q^n \in M_k(\mathbb{Q}_p)$, and if $k \not\equiv 0 \pmod{p^{m-1}(p-1)}$ then

$$v_p(a_0) + m \geq \inf_{n\geq 1} v_p(a_n)$$

6.2 Another Inverse Limit: Gal($\overline{\mathbb{Q}}/\mathbb{Q}$)

Let $\overline{\mathbb{Q}}$ be the algebraic closure of $\mathbb{Q}$ in $\mathbb{C}$. Then every element $\alpha \in \overline{\mathbb{Q}}$ lies in some finite Galois extension of $\mathbb{Q}$ (for example, the splitting field of the minimal polynomial of α).

If $L/\mathbb{Q}$ is a finite Galois extension (with $L \subset \overline{\mathbb{Q}}$), then any automorphism $\sigma : \overline{\mathbb{Q}} \to \overline{\mathbb{Q}}$ which acts as the identity on $\mathbb{Q}$ restricts to an automorphism $\overline{\sigma} : L \to L$ over $\mathbb{Q}$ (this can be viewed as a defining property of being Galois). So we get a map

$$\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{Gal}(L/\mathbb{Q})$$

which is surjective (by a Zorn's lemma argument).

We want to study $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. Let $\{L_i\}_{i\in I}$ be the finite Galois extensions of $\mathbb{Q}$ in $\overline{\mathbb{Q}}$, and let $G_i = \mathrm{Gal}(L_i/\mathbb{Q})$. If $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ and $\alpha \in \overline{\mathbb{Q}}$, then $\alpha \in L_i$ for some $L_i/\mathbb{Q}$ finite, and $\sigma(\alpha) = (\pi_i\sigma)(\alpha)$, so σ **is determined by its images in** G_i. Hence

$$G_{\mathbb{Q}} := \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \cong \varprojlim_{L_i/\mathbb{Q}\text{ fin.Gal.}} \mathrm{Gal}(L/\mathbb{Q})$$

This inverse limit construction is analogous to the construction of $\mathbb{Z}_p$ as $\varprojlim(\mathbb{Z}/p^n\mathbb{Z})$, except that here the system $\{L_i\}$ of finite Galois extensions is not linearly ordered. Still, for any two finite Galois extensions L_i and L_j of $\mathbb{Q}$, there exists another finite Galois extension L_{ij} containing them both, allowing us to talk about compatibility in the inverse limit.

Just as in the case of the p-adics, we can define a topology on the inverse limit. In this case, we define the topology abstractly with no reference to a metric: it is the coarsest topology such that every projection $G_{\mathbb{Q}} \to \mathrm{Gal}(L, Q)$, where L is finite Galois over $\mathbb{Q}$, is continuous, where $\mathrm{Gal}(L, Q)$ is equipped with the discrete topology. This is called the *profinite topology*. To gain intuition about this construction, think about why the same construction for $\mathbb{Z}_p$ as the inverse limit $\lim\limits_{\leftarrow}(\mathbb{Z}/p^n\mathbb{Z})$ recovers the topology on $\mathbb{Z}_p$.

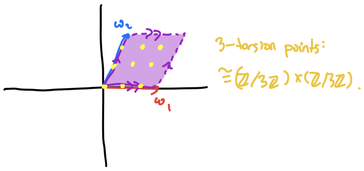
6.3 Galois Representations

6.3.1 Elliptic Curve Galois Representations

Let E be an elliptic curve defined over $\mathbb{Q}$. Recall that the complex points of E have an abelian group structure, where the point at infinity is the identity, and that the group of n-torsion points of E is defined as

$$E[n] := \{P : P \in E(\mathbb{C}), n \cdot P = 0\}$$

and we have $E[n] \cong (\mathbb{Z}/n\mathbb{Z})^2$.


ω_2
ω_1
3-torsion points:
$\cong (\mathbb{Z}/3\mathbb{Z}) \times (\mathbb{Z}/3\mathbb{Z})$.

Moreover, these torsion points have coefficients in $\overline{\mathbb{Q}}$ since being torsion can be expressed as an algebraic condition on each coordinate involving the coefficients of the equation defining E. So the Galois group $G_{\mathbb{Q}}$ acts on the torsion points. In particular we will be interested in its action on $E[p^n] \cong (\mathbb{Z}/p^n\mathbb{Z})^2$ for a prime p. These compatible actions taken in the limit as $n \to \infty$ induce an action of $G_{\mathbb{Q}}$ on $(\mathbb{Z}_p)^2$. Each $\sigma \in G_{\mathbb{Q}}$ acts on $(\mathbb{Z}_p)^2$ linearly, so we get a group homomorphism

$$\rho_E : G_{\mathbb{Q}} \to GL_2(\mathbb{Z}_p)$$

In fact, this homomorphism is *continuous* (both groups are equipped with a topology). This motivates the following definition:

Definition 6.5

A **two-dimensional Galois representation over** A is a continuous homomorphism

$$\rho : G_{\mathbb{Q}} \to GL_2(A)$$

We can also associate to ρ a determinant function $\det \rho : G_{\mathbb{Q}} \to A^\times$ as the composition

$$G_{\mathbb{Q}} \xrightarrow{\rho} GL_2(A) \xrightarrow{\det} A^\times$$

6.3.2 Local Galois Groups

Let ℓ be a prime and let

$$G_\ell := \mathrm{Gal}(\overline{\mathbb{Q}_\ell}/\mathbb{Q}_\ell)$$
$$G_\infty := \mathrm{Gal}(\mathbb{C}/\mathbb{R})$$

Also, any $\sigma \in G_\ell$ restricts to an automorphism of $\overline{\mathbb{Q}}$ over $\mathbb{Q}$ since $\overline{\mathbb{Q}}$ is Galois over $\mathbb{Q}$. And since $\overline{Q}$ is dense in $\overline{\mathbb{Q}_\ell}$, the restriction is injective, so

$$G_\ell \subseteq G_{\mathbb{Q}}.$$

6.3.3 Inertia and Ramification

We can define a subgroup of G_ℓ of automorphisms with specified behavior at the point ℓ.

For $\ell \neq \infty$, let

$$\overline{\mathbb{Z}}_\ell := \{x \in \overline{\mathbb{Q}}_\ell : |x|_\ell \leq 1\}$$
$$\lambda := \{x \in \overline{\mathbb{Q}}_\ell : |x|_\ell < 1\}$$

Then $\overline{\mathbb{Z}}_\ell/\lambda \cong \overline{\mathbb{F}}_\ell$ and G_ℓ acts on $\overline{\mathbb{F}}_\ell$.

Definition 6.6

The inertia group at ℓ is

$$I_\ell := \{\sigma \in G_\ell : \sigma \text{ acts as the identity on } \overline{\mathbb{F}_\ell}\}$$

Now we can look into "local properties" of Galois representations. For a ring A and a global Galois representation $\rho : G_{\mathbb{Q}} \to GL_2(A)$, restrictions give local Galois reprsentations

$$\rho|_{G_\ell} G_\ell \to GL_2(A)$$

Definition 6.7

Properties of a representation $\rho : G_{\mathbb{Q}} \to GL_2(A)$

1. ρ is **odd** if, for c the element of $\mathrm{Gal}(\mathbb{C}/\mathbb{R})$ corresponding to complex conjugation, $\det \rho(c) = -1$.

2. for a prime ℓ, ρ is **unramified at** ℓ if $I_\ell \subseteq \ker \rho|_{G_\ell}$.

3. ρ is **flat** at a prime p if $\forall$ ideals $J \subset A$ such that A/J is finite, then $\overline{\rho} : G_p \to GL_2(A/J)$ extends to a finite flat group scheme.

4. ρ is **irreducible** if it has no nontrivial subrepresentation.

6.4 Fermat’s Last Theorem

Theorem 6.8

The equation $a^n + b^n = c^n$ has no nontrivial integer solutions if $n \geq 3$.

““““‘Proof”””””””””””: If $n = pm$ for a prime p, then the equation can be rewritten as

$$(a^m)^p + (b^m)^p = (c^m)^p$$

so it suffices to prove FLT for $n = p$ and $n = 4$ (but the case $n = 4$ was done by Fermat).

Suppose for contradiction that $a^p + b^p = c^p$ for a, b, c coprime integers such that $abc \neq 0$. We can associate to this curve an elliptic curve, the Frey curve

$$E_{a^p,b^p,c^p} : y^2 = x(x - a^p)(x + b^p)$$

The Galois representation associated to this has some remarkable properties.

Theorem 6.9 Frey, Serre

Let $p \geq 5$ prime and $a, b, c \in \mathbb{Z}$ such that $a^p + b^p + c^p = 0$. Suppose $a \equiv -1 \pmod 4$ and $2|b$. Then $\overline{\rho}_{a^p,b^p,c^p}$ is absolute irreducible, odd, and unramified outside of $2, p$ and flat at p.

This theorem will be a black box for us.

These properties are so remarkable that people suspect that **no** Galois representation has them.

We try to get at $\overline{\rho}_{a^p,b^p,c^p}$ another way, via modular forms.

The Eichler–Shimura construction gives us a way of associated an elliptic curve to a level N rational newform $f \to E_f$ such that the conductor N of E_f equals the level of f, and the traces of Frobenius $a_p(E)$ encode coefficients of f.

The modularity theorem (Wiles, Taylor–Wiles, Breuil–Conrad–Diamond–Taylor) allows us to conclude that every elliptic curve arises this way! So we can associate to E_{a^p,b^p,c^p} a modular form, and use results about Galois representations associated to modular forms to derive a contradiction.

Theorem 6.10 Ribet

Let f be a weight 2 newform of level N, and let ℓ be a prime such that $\ell \mid N$ but $\ell^2 \nmid N$. Suppose $\overline{\rho}_f$ is absolutely irreducible and that $\overline{\rho}_f$ is unramified at ℓ or $\ell = p$ and $\overline{\rho}_f$ is flat at p.
Then there is a weight 2 newform of level N/ℓ such that $\overline{\rho}_f \cong \overline{\rho}_g$.

We now have all the pieces we need to outline a proof of FLT. We note that $p = 3$ had already been completed by Euler and $n = 4$ completed by Fermat. First, E_{a^p,b^p,c^p} has an associated modular form f_{a^p,b^p,c^p} by modularity theorem. And $\overline{\rho}_{a^p,b^p,c^p}$ is "barely ramified" etc by Frey–Serre. So we can apply Ribet's Theorem iteratively to the primes dividing the conductor $N = \prod_{\ell \mid abc} \ell$ since N is squarefree; this procedure produces a newform g of weight 2 and level 2. But the space of level 2 weight 2 cusp forms $S_2(\Gamma_0(2))$ has dimension equal to the genus of $X_0(2)$ which is 0 (see Dr. Watson's lectures). So there is no such form g, a contradiction!

So we've established the veracity of Fermat's Last Theorem. But what's the big deal? Why do we care so much about FLT, what are its applications? I close with a corollary that would have shocked the Pythagoreans. And I'm making this choice precisely because there's no risk of any of you throwing me off a boat for providing such a ridiculous proof.

Corollary 6.11

$\sqrt[3]{2}$ is irrational.

Proof: Let $c = \sqrt[3]{2}$. If c were in $\mathbb{Q}$, then $1^3 + 1^3 = c^3$ would be a rational solution to $a^3 + b^3 = c^3$. But there is no such solution by Fermat's Last Theorem. □

I am just joking. The real reason we care about FLT is all of the cool math we got to learn in order to (begin to) understand the proof! The real treasure is the friends we made along the way

The real reason we care about FLT is the friends we made along the way!