Enumerating mod p eigenforms

Craig Citro
(joint with Alex Ghitza)

October 31, 2008

Outline

1. Modular forms for the uninitiated

2. Modular forms for the initiated

3. Modular forms mod p

4. Galois representations

5. Finiteness Theorems and Results

6. Practice

Goals & Status

Goal

Enumerate all systems of eigenvalues coming from mod p modular forms.

Current Status

We can currently do this for level 1.

Outline

1 Modular forms for the uninitiated

2 Modular forms for the initiated

3 Modular forms mod p

4 Galois representations

5 Finiteness Theorems and Results

6 Practice

Modular Forms (for the uninitiated)

A **modular form** is a holomorphic function on the upper half plane $\mathfrak{H}$ that satisfies a strange-looking symmetry condition and a growth condition. For any modular form f, we have the following invariants:

- a **weight**, which is a positive integer (and is "usually" even), always denoted k
- a **level**, which is a positive integer, always denoted N
- a **Fourier expansion**, or q-**expansion**, which is a power series expansion for f in $q = e^{2\pi iz}$ of the form

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

If $a_0 = 0$, we call f a **cusp form**. Given integers N and k, we write $M_k(N)$ for the modular forms with weight k and level N, and $S_k(N)$ for the space of cusp forms.

Modular Forms (for the uninitiated)

Fact

For any positive integers k and N, $M_k(N)$ is a finite-dimensional vector space over $\mathbb{C}$.

Fact

A modular form over $\mathbb{C}$ of level N is determined by its q-expansion.

Fact

Let E be an elliptic curve over $\mathbb{Q}$ with conductor N, and write $L(s, E) = \sum_n a_n n^{-s}$ for its Hasse-Weil L-function. Then $\sum_n a_n q^n$ is a cusp form in $S_2(N)$.

Fact

There is a commuting family of operators on $M_k(N)$ and $S_k(N)$ called **Hecke operators**, *one for each integer n, which we denote T_n or $T_{n,k}$. We write $\mathscr{H}$ or $\mathscr{H}_N$ for the algebra $\mathbb{Z}[T_n|(N,n)=1]$. Much of the arithmetic theory of modular forms is centered around these operators, and in particular, the forms which are simultaneous eigenvectors for all these operators.*

Modular Forms (for the uninitiated)

For completeness, I'll also write down a definition. Let $\Gamma_0(N)$ be the subgroup of $\mathrm{SL}(2,\mathbb{Z})$ (that is, the 2×2 matrices with integer entries and determinant 1) whose lower-left entry is divisible by N. Then a **modular form of weight k and level $\Gamma_0(N)$** is a holomorphic function $f : \mathfrak{H} \to \mathbb{C}$ satisfying:

- $f\left(\frac{az+b}{cz+d}\right) = (cz+d)^k f(z)$ for all $z \in \mathfrak{H}$, $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N)$.
- The sequence of coefficients $\{a_n\}$ in the q-expansion of f is $O(n^r)$ for some r.

Outline

1. Modular forms for the uninitiated
2. Modular forms for the initiated
3. Modular forms mod p
4. Galois representations
5. Finiteness Theorems and Results
6. Practice

Modular Forms over $\mathbb{C}$ (for the initiated)

- $M_k(\Gamma) =$ modular forms of weight k, level Γ over $\mathbb{C}$
- $S_k(\Gamma) =$ cuspidal subspace of $M_k(\Gamma)$
- $M_k(\Gamma, R) =$ modular forms in $M_k(\Gamma)$ whose Fourier expansions have coefficients in R

Outline

1. Modular forms for the uninitiated

2. Modular forms for the initiated

3. Modular forms mod p

4. Galois representations

5. Finiteness Theorems and Results

6. Practice

Modular forms mod p

We now want to study modular forms mod p. However, our definition of a modular form as a holomorphic function is too analytic to be helpful here. However, we noted that a modular form is determined by its q-expansion, which we **can** think about trying to reduce mod p, at least when the coefficients are in $\mathbb{Z}$ (or more generally in $\mathbb{Z}_{(p)}$, i.e. those rational numbers whose denominator is coprime to p).

Fact

Given any k and N, there exist forms $f_1, \ldots, f_d \in M_k(N, \mathbb{Z})$ whose $\mathbb{C}$-span is $M_k(N, \mathbb{C})$.

Then we have natural maps

$$M_k(N, \mathbb{Z}) \hookrightarrow \mathbb{Z}[[q]] \to \mathbb{F}_p[[q]].$$

We now define the space of **mod p modular forms** to be the $\mathbb{F}_p$-span of the image of this composite map. We abuse notation and denote this $M_k(N, \mathbb{F}_p)$, or (even worse) $M_k(N)$ when the context is clear.

Modular forms mod p

Of course, this definition is completely ad-hoc, and raises several questions:

1. Is every element of $M_k(N, \mathbb{F}_p)$ the reduction mod p of an element of $M_k(N, \mathbb{Z})$?
2. Does the space get larger if we add in the reductions of elements in $\mathbb{Z}_{(p)}$?
3. What kind of bases for $M_k(N, \mathbb{Z})$ generate all of $M_k(N, \mathbb{F}_p)$?
4. What relationship do $\dim_{\mathbb{C}} M_k(N, \mathbb{C})$ and $\dim_{\mathbb{F}_p} M_k(N, \mathbb{F}_p)$ have?

Of course, there are better definitions. First, one can take the $\mathbb{F}_p$-span of the entire space $M_k(N, \mathbb{Z}_{(p)})$. The "best" choice is to have a completely geometric definition, with modular forms global sections of a certain line bundle on a modular curve, which is the solution of a certain moduli problem. In this setting, the geometric theory of mod p and p-adic modular forms is due to Katz. In some cases (i.e. when N or p is less than 5), one has to be careful with the geometric definition, but it's possible to get all the details squared away.

Modular forms mod p: What's the same?

Many things are fairly similar to the characteristic 0 theory. In particular, you still have q-expansions (which is how we defined them, but the q-expansion map still makes sense with the geometric definition), and Hecke operators. We write

$$\mathscr{H} = \mathbb{Z}[T_\ell \mid \ell \nmid Np]$$

for the Hecke algebra acting on mod p modular forms of level N. (One should really think of $\mathscr{H}$ as its own object, and we're just talking about its image in $\mathrm{End}(M_k(N, \mathbb{F}_p))$.) One also has a notion of weight and level, as the notation suggests.

However, things are not all as they seem ...

Modular forms mod p: What's different?

Consider the Eisenstein series E_{p-1}, normalized to have constant coefficient 1. Then the Fourier expansion is given by

$$E_{p-1}(z) = 1 - \frac{2(p-1)}{B_{p-1}} \sum_{n \geq 1} \sigma_{p-2}(n) q^n.$$

However, the theorem of Kummer and von Staudt says that B_{p-1} has denominator divisible by p. Thus the reduction of this form mod p has q-expansion 1.

That is, we have a form of weight other than 0 whose q-expansion is just 1. (For the geometrically inclined, this form is simply the mod p reduction of the Hasse invariant.) We'll also denote this form A.

Modular forms mod p: What's different?

So this means that we can't identify a modular form of a given level simply by looking at its q-expansion: we need to know something about its weight. But don't despair! In fact, the existence of A makes so many new things possible.

Let $M(N) = \bigoplus_k M_k(N)$ denote the **ring of mod p modular forms**.

Modular forms mod p: What's different?

As with the case of modular forms in characteristic 0, for any given space $M_k(N)$, we have a q-expansion map

$$q : M_k(N) \hookrightarrow \mathbb{F}_p[[q]].$$

However, we know that forms of different weights different weights can have the same q expansion. That is, the map

$$q : M(N) \longrightarrow \mathbb{F}_p[[q]]$$

is **not** an injection. However, all is not lost: one knows that the kernel of the above map is given by $A - 1$, where A is the mod p reduction of the Hasse invariant, which is a modular form of weight $p - 1$. In particular, if two forms have the same q-expansion, then we know that their weights are congruent mod $p - 1$.

Modular forms mod p: What's different?

We said that the central objects for number theorists are the **eigenforms**, i.e. simultaneous eigenvectors for all the Hecke operators. Then a simultaneous eigenform f for $\mathscr{H}$ gives rise to a system of eigenvalues for each T_ℓ, which we'll call a **mod p eigensystem** for short. (We should probably include N in the notation, but we're currently focused on level 1, so we won't here.) This is the same as a homomorphism

$$\lambda_f : \mathscr{H} \to \overline{\mathbb{F}_p}.$$

We'll also write $M(N)$ for the ring of all mod p modular forms of level N, i.e. $M(N) = \oplus_k M_k(N)$.

Mod p modular forms: What's different?

One also has an operator

$$\theta : M_k(N) \to M_{k+p+1}(N),$$

whose effect on q-expansions is just $q \cdot d/dq$, and satisfies the relation

$$T_\ell(\theta f) = \ell \theta T_\ell(f).$$

In particular, if f is an eigenform with eigenvalues $\{a_\ell\}$, then θf is an eigenform with eigenvalues $\{\ell a_\ell\}$.

Outline

1. Modular forms for the uninitiated
2. Modular forms for the initiated
3. Modular forms mod p
4. **Galois representations**
5. Finiteness Theorems and Results
6. Practice

Galois representations

There is another interesting interpretation for these mod p eigensystems. Let

$$\rho : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}(2, \overline{\mathbb{F}_p})$$

be an odd continuous galois representation of conductor N. (**Odd** just means that $\det(\rho(c)) = -1$ for any complex conjugation c.) Then ρ is determined (at least up to semisimplification) by the collection of $\{a_\ell\}_{\ell \neq p}$, where

$$a_\ell = \mathrm{tr}(\rho(\mathrm{Frob}_\ell)).$$

Serre's Conjecture

Any system of $\{a_\ell\}$ coming from such a representation in fact arises from a mod p modular form of level N and some (predictable) weight k.

This is now a Theorem of Khare, Wintenberger, Kisin and others.

Galois representations

In particular, the θ operator has a meaningful interpretation in terms of galois representations. Write ε_p for the mod p cyclotomic character, i.e. the character

$$\varepsilon_p : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathbb{F}_p^\times$$

given by

$$\sigma(\zeta_p) = \zeta_p^{\varepsilon_p(\sigma)},$$

where ζ_p is a primitive p-th root of unity. Then one knows that $\varepsilon_p(\mathrm{Frob}_\ell) = \ell \in \mathbb{F}_p^\times$ for all $(\ell, p) = 1$. So if f is a mod p eigenform with associated mod p galois representation ρ_f, we have

$$\rho_{\theta f} = \rho_f \otimes \varepsilon_p.$$

Outline

1. Modular forms for the uninitiated
2. Modular forms for the initiated
3. Modular forms mod p
4. Galois representations
5. Finiteness Theorems and Results
6. Practice

Finiteness Theorems, I

First, we need to know that it's possible to enumerate **all** eigensystems for a given N:

Theorem (Serre, Tate, Jochnowitz)

There are only finitely many eigensystems arising from $M(N)$. In particular, all eigensystems occur in weight at most $p+1$ up to twist.

In particular, by applying θ to the space of weight 1 forms, it suffices to look in weights up to $p+2$ to find all eigensystems.

Finiteness Theorems, II

Now one needs a way of comparing two eigensystems in a finite amount of time:

Theorem (Sturm)

Let $f, g \in M_k(N)$. If $a_n(f) = a_n(g)$ for all $n \leq k[\Gamma(1) : \Gamma_0(N)]/12$. Then $f = g$.

We call $\mathrm{Sturm}(k)$ the bound occurring in this theorem for weight k.

Theorem (C., Ghitza)

Let $f \in M_{k_1}(1)$, $g \in M_{k_2}(1)$. Say that $k_1 \equiv k_2 \mod p-1$, and let k' be the larger of the residues of k_1 and k_2 mod $p+1$. If $a_n(f) = a_n(g)$ for $n \leq \mathrm{Sturm}(k'+p+1)$, then $f = g$.

Outline

1 Modular forms for the uninitiated

2 Modular forms for the initiated

3 Modular forms mod p

4 Galois representations

5 Finiteness Theorems and Results

6 Practice

We currently have two implementations for enumerating mod p eigensystems, via completely different means:

- using modular symbols
- working with q-expansions

(Both use Sage.)

Modular Symbols Implementation

This works by creating each space of mod p modular symbols, and simply computing the decomposition of the space and asking for the eigenvalues.
Pros:

- Works for all levels
- Easier to write

Cons:

- Vastly slower
- Mod p modular symbols are more subtle than in the case of characteristic zero.

q-expansion Implementation

The second implementation simply computes a basis of q-expansions for $M_k(\mathrm{SL}(2, \mathbb{Z}))$ in characteristic 0 (specifically the Victor Miller basis), reduces mod p, and uses some clever linear algebra techniques to find eigenforms for the Hecke operators one at a time.
Pros:

- Quite fast
- Less subtle than mod p modular symbols
- Always does **as little** linear algebra as possible

Cons:

- Only currently works for level 1

Demo!

Here’s a live demo of the code running ...

Thanks for listening!

Any questions?