Approaches to computing overconvergent p-adic modular forms

David Loeffler (Imperial College, London)

Computations with Modular Forms workshop, 18th August 2008

1 Introduction to overconvergent modular forms

1.1 Motivation

Let p be a prime number. It's well known that classical modular Hecke eigenforms can satisfy nontrivial congruence relations modulo powers of p; for example the standard Eisenstein series E_k satisfies $E_{p-1} = 1 \pmod p$, and more generally $E_{(p-1)p^r} = 1 \pmod{p^{r+1}}$ for all integers $r \geq 0$. We'd like to construct some kind of p-adic space of modular forms in which this really represents a limiting process, with $E_{(p-1)p^r}$ being a sequence of modular forms tending towards the constant form 1 in weight 0.

An early attempt to construct such a space dates back to work of Serre in the early 1970's [Ser73]. Serre considered the space of formal q-expansions that are uniform p-adic limits of q-expansions of modular forms of fixed level N and arbitrary weight. This space, the space of *convergent p-adic modular forms*, is infinite-dimensional but nonetheless has many good properties; for example, such forms have a well-defined weight, and one can define an action of Hecke operators by the usual formulae on q-expansions and these turn out to be continuous. However, this space has too many eigenforms: a construction due to Gouvea shows that one may construct eigenfunctions for the U_p operator with arbitrary eigenvalue, so in particular these p-adic eigenfunctions contain no interesting arithmetical information. We shall see in the remainder of this section how to construct a space which remedies this, and it is the elements of this space which we will attempt to compute in the following sections.

1.2 Geometry of modular curves

We start by recalling some standard facts from the algebro-geometric theory of modular forms. (An excellent summary of these constructions can be found in the appendix to Katz's article [Kat73].) Let $\Gamma \subseteq \mathrm{SL}_2(\mathbb{Z})$ be a congruence subgroup. Then the quotient $Y(\Gamma) = \mathcal{H}/\Gamma$ can be identified with the complex points of an algebraic curve. We shall mostly be interested in $\Gamma = \Gamma_0(Np)$ or $\Gamma_0(N)$, or the corresponding Γ_1 versions. For these Γ, if N is sufficiently large ($N \geq 5$ is sufficient), then the algebraic curve $Y(\Gamma)$ has a well-understood canonical model over $\mathbb{Z}$, which is a fine moduli space classifying elliptic curves with suitable level structure; and it posesses a canonical compactification $X(\Gamma)$ which is a proper scheme over $\mathbb{Z}$. By considering differentials on the universal elliptic curve over $Y(\Gamma)$, one can construct a line bundle ω on $X(\Gamma)$ with the convenient property that $H^0(X(\Gamma), \omega^{\otimes k}) = M_k(\Gamma, \mathbb{Q})$. (*Remark*: The requirement that the level be ≥ 5 can be worked around, by the usual expedient of adding in auxilliary level structure and then taking invariants under it.)

It's well known that if X is a projective curve over $\mathbb{C}$, then algebraic sections and holomorphic sections of line bundles over X coincide; this is Serre's "GAGA" principle. But we are interested in p-adic analysis here, so instead we consider X as a projective curve over $\mathbb{Q}_p$, to which we can associate an object known as a rigid analytic space – this is a kind of p-adic analogue of a manifold, with functions being given by p-adically convergent power series. There is a GAGA principle for rigid spaces due to Köpf ([Köp74]), and we find that $H^0_{rig}(X(\Gamma)_{\mathbb{Q}_p}, \omega^{\otimes k}) = M_k(\Gamma, \mathbb{Q}) \otimes_{\mathbb{Q}} \mathbb{Q}_p$; so we have in a sense placed our space in a "p-adic context".

1.3 Butchered modular curves

An insight due to Katz is that the "bad points" of $X(\Gamma)$ are those corresponding to elliptic curves with supersingular reduction modulo p; if the congruence of q-expansions $E_{p-1} = 1 \pmod p$ is to be given a geometric interpretation, then points where E_{p-1} vanishes modulo p are troublesome.

In the rigid space associated to $X(\Gamma)$, these supersingular points are a well-behaved rigid subspace, isomorphic to a finite union of open p-adic discs corresponding to the supersingular j-invariants in characteristic p. Hence one can construct the complement of these discs as a rigid space, which is denoted $X(\Gamma)^{ord}$ (the *ordinary locus*).

Theorem (Katz). *The space $H^0_{rig}(X(\Gamma)^{ord}, \omega^{\otimes k})$ is isomorphic (as a Banach space and as a Hecke module) to Serre's space of convergent p-adic modular forms.*

This shows that we are on the right track, but we have been a little too drastic in throwing away the whole discs – we have found ourselves in a space that is too large. So we try throwing away smaller subdiscs; but one must do this in some reasonably canonical manner (so that the resulting space is stable under the correspondences on $X(\Gamma)$ that give Hecke operators). So we run in to the usual problem that discs in ultrametric spaces don't know where their centres are.

Fortunately, there is a solution. The mod p Hasse invariant, a modular form over $\mathbb{F}_p$ of weight $p-1$ which is zero precisely at supersingular points, can be lifted to a characteristic 0 form defined over $\mathbb{Q}_p$; for example, if $p \geq 5$ the level 1 Eisenstein series is such a lift. Katz shows that if A is such a lifting, and x is a point of $X(\Gamma)(\mathbb{Q}_p)$ such that $\operatorname{ord}_p A(x) < 1$, then this number $\operatorname{ord}_p A(x)$ is independent of the choice of A. So if $0 < r < 1$, defining $X(\Gamma)^{\leq r}$ to be the set of x such that $\operatorname{ord}_p A(x) \leq r$, we obtain a well-defined rigid subspace of $X(\Gamma)$. Katz moreover shows that if $r < \frac{p}{p+1}$ this space is stable under the Hecke correspondences; hence the space $H^0_{rig}(X(\Gamma)^{\leq r}, \omega^{\otimes k})$ gives a space with a continuous Hecke action. This is the space of r-overconvergent p-adic modular forms, denoted $M_k^\dagger(\Gamma, r)$. Finally, on this space the Hecke operator U_p is compact – in fact this follows from the fact that it increases overconvergence, so if f is r-overconvergent $U_p f$ will be pr-overconvergent – so we have a good spectral theory. This is the space we shall attempt to compute.

(A more detailed introduction to this theory can be found in Emerton's survey article [Eme].)

2 Explicit parametrisation

The first algorithm we shall discuss is based on work of several authors including Emerton, Smithline, Buzzard and Calegari; a detailed description and some tables of results can be found in my paper [Loe07].

If $\Gamma = \mathrm{SL}_2(\mathbb{Z})$, then $X(\Gamma)$ is isomorphic to $\mathbb{P}^1$. If p is sufficiently small, then there is only one supersingular j-invariant, so for any r, $X(\Gamma)^{\leq r}$ will be the complement in $\mathbb{P}^1$ of an open disc, which is simply a closed disc. So the space of functions on $X(\Gamma)^{\leq r}$ can be written down explicitly. This works for p in the set $\{2, 3, 5, 7, 13\}$.

We shall define an explicit uniformiser as follows. Let f be the function $\left(\frac{\Delta(pz)}{\Delta(z)}\right)^{1/(p-1)}$. This is a function on $X_0(p)$; but Katz's theory of the canonical subgroup shows that for $r < \frac{p}{p+1}$, the forgetful map $X_0(p) \to X_0(1)$ has a rigid-analytic section over $X_0(1)^{\leq r}$, identifying it with $X_0(p)^{\leq r}$. One can show explicitly that $X_0(p)^{\leq r}$ is precisely the region of $X_0(p)$ where $\operatorname{ord}_p f \geq \frac{-12r}{p-1}$.

Hence the space of rigid functions on $X_0(1)^{\leq r}$ is the space of power series

$$\Big\{ \sum_{n\geq 0} a_n f^n \;\Big|\; a_n \in \mathbb{Q}_p, \operatorname{ord}_p(a_n) - \frac{12rn}{p-1} \to +\infty \Big\}.$$

This space is, by definition, $M_0^\dagger(\mathrm{SL}_2(\mathbb{Z}), r)$; and if E_k^* represents the level p Eisenstein series of weight k whose U_p-eigenvalue is 1, then (by a theorem of Coleman) E_k^* is overconvergent with no zeros on $X_0(p)^{\leq r}$, and hence $M_k^\dagger(\mathrm{SL}_2(\mathbb{Z}), r)$ is simply $E_k^* \cdot M_0^\dagger(\mathrm{SL}_2(\mathbb{Z}), r)$. (In particular, $M_k^\dagger$ and $M_0^\dagger$ are isomorphic as abstract Banach spaces; but the Hecke actions on the two are different.)

So we would like to compute the matrix of the U_p operator in this basis; equivalently, we want to write $(E_k^*)^{-1} \cdot U_p(f^j E_k^*)$ as a convergent power series in f. In principle we can calculate this directly from the q-expansion – since the q-expansion of f has no constant term, the first n coefficients of the q-expansion of $U_p(f^j E_k^*)$ determine the first n entries of the corresponding column of the matrix. However, computing many terms of the matrix this way is prohibitively expensive.

A more practical method relies on a cunning observation due originally to an obscure Norwegian author named Kolberg [Kol61]. In terms of q-expansions,

$$U_p(f^j E_k^*)(q) = (f^j E_k^*)(q^{1/p}) + (f^j E_k^*)(\zeta_p \cdot q^{1/p}) + \cdots + (f^j E_k^*)(\zeta_p^{p-1} \cdot q^{1/p});$$

so if k and q are fixed, $U_p(f_j E_k^*)$ will satisfy a recurrence of degree p in j, with coefficients that are the elementary symmetric functions of the set $\{f(q^{1/p}), \ldots, f(\zeta_p^{p-1} \cdot q^{1/p})\}$.

Lemma (Kolberg). *These are all polynomials in f of degree at most p.*

This motivates the choice of function f. (For a "random" uniformiser, one would expect them to be rational functions of f, but the fact that f is an η-product constrains the denominators to be trivial.) For example, if $p = 2$ then we find that $U_p(f^j E_k^*) = (48f + 2^{11} f^2) U_p(f^{j-1} E_k^*) + f U_p(f^{j-2} E_k^*)$.

This recurrence allows the entire matrix of U_p on $M_k^\dagger(\Gamma, r)$, for any given r and k, to be computed very rapidly using only integer arithmetic, given only its first p rows and columns. Since this matrix represents a compact operator for r in the appropriate range, we may truncate it to an $N \times N$ top left submatrix and calculate the eigenvalues and eigenvectors of this truncated matrix; these will converge rapidly to the eigenvectors of U_p with small slope (p-adic valuation of the eigenvalue). Converting the f-adic expansions of these back to q-expansions it is also easy to read off the action of the Hecke operators at other primes $\ell \neq p$, so one has an essentially complete description of $M_k^\dagger$.

(*Remark:* If p is 11, 17 or 19, then there are two supersingular j-invariants and a similar idea can be made to work: the overconvergent locus is no longer a disc but an annulus, so one is forced to use Laurent series in place of power series. However, there is no analogue of the uniformiser f and the associated recurrence formula, so calculating the Hecke action is harder, but still possible; this has been implemented by Lloyd Kilford. For primes ≥ 23 it is not clear how to proceed.)

3 The p-adic trace formula

The second algorithm we shall use, which I learnt from Frank Calegari, gives less information but in a much wider range of cases.

There is a classical formula, due to Eichler, which gives the trace of the Hecke operator T_m acting on $S_k(\Gamma_0(N))$ (originally assuming N is square-free). This was subsequently generalised vastly by Selberg and others to give a whole industry of trace formulae for many different kinds of automorphic forms; but we shall use only Eichler's original formula, and a generalisation by Hijikata to handle the case where N is not square-free.

The dependency on m and N is subtle, involving a sum over factorisations of m in all possible imaginary quadratic fields, but for a fixed m and N it depends on k only through a finite sum of terms of the form ρ^{k-1}, where ρ is an algebraic integer constant. For example, $\mathrm{Tr}(T_2 | S_k(\mathrm{SL}_2 \mathbb{Z}))$ is given by

$$-\frac{(1+\sqrt{-1})^{k-1} - (1-\sqrt{-1})^{k-1}}{4\sqrt{-1}} - \frac{\left(\frac{1+\sqrt{-7}}{2}\right)^{k-1} - \left(\frac{1-\sqrt{-7}}{2}\right)^{k-1}}{\sqrt{-7}} - \frac{\sqrt{-2}^{k-1}}{2} - 1.$$

Let p be a fixed prime, and N some arbitrary level with $p \nmid N$. Fix an integer k_∞ and consider a cunningly chosen sequence of integers k_n satisfying the following conditions:

- $k_n = k_\infty \pmod{p-1}$ for all n (or mod 2 if $p = 2$)
- $k_n \to \infty$ in $\mathbb{R}$
- $k_n \to k_\infty$ in $\mathbb{Z}_p$

In the trace formula for $\mathrm{Tr}(T_p | S_k(\Gamma_0(N)))$, whenever there is a term ρ^{k-1} with the p-adic valuation of ρ positive, $\rho^{k_n - 1}$ will tend rapidly to zero p-adically, since k_n is growing large in the archimedean sense. However, if ρ is a p-adic unit, then the first and last requirements imply that $\rho^{k_n - 1}$ converges to $\rho^{k_\infty - 1}$. So

$$\lim_{n\to\infty} \mathrm{Tr}(T_p | S_{k_n}(\Gamma_0(N)))$$

can be calculated by simply deleting all terms in the trace formula with ρ not a p-adic unit, and evaluating the remaining terms at k_∞.

Theorem. *This limit is equal to the trace of U_p acting on $S^\dagger_{k_\infty}(\Gamma_0(N), r)$, for all sufficiently small $r > 0$.*

This is a consequence of two deep theorems of Coleman: firstly, U_p-eigenvalues vary continuously as the weight varies p-adically, so for n large the eigenvalues of U_p on $S^\dagger_{k_n}$ are close to those on $S^\dagger_{k_\infty}$. Secondly, if f is a U_p-eigenform in $S^\dagger_k(\Gamma_0(N))$ for integer k and the valuation of its U_p-eigenvalue is $< k - 1$, then f is in fact a classical modular form of level $\Gamma_0(Np)$. Hence the trace of U_p acting on overconvergent forms of level p differs from the trace on classical forms by an error term of order at most p^{k-1}. Finally, the difference between the trace of U_p in level $\Gamma_0(N)$ and $\Gamma_0(Np)$ is of the order of $p^{(k-2)/2}$. Combining all of these facts, we see that if k_n is sufficiently large in the archimedean sense, and sufficiently close to k_∞ in the p-adic sense, $\mathrm{Tr}(T_p|S_{k_n})$ is close to $\mathrm{Tr}(U_p|S^\dagger_{k_\infty})$.

Extensions to the method

The proof of the above result can be generalised: one does not need to take the limit of T_m for $m = p$, but merely for some m with $p|m$. For instance, one can take $m = p^e$ and deduce that

$$\lim_{n\to\infty} \mathrm{Tr}(T_{p^e} | S_{k_n}(\Gamma_0(N))) = \mathrm{Tr}((U_p)^e | S_{k_\infty}(\Gamma_0(N))),$$

and computing this for $e = 1 \dots r$ and applying a simple identity one obtains the first r coefficients of the characteristic power series $\det(1 - tU_p)$, from which one can derive approximations to the U_p-eigenvalues.

In principle one can even consider the limit of $\mathrm{Tr}(T_{p^e \ell})$ where ℓ is a prime not dividing pN; this gives the trace of $(U_p)^e \cdot T_\ell$, and by linear algebra one can recover the T_ℓ-eigenvalues corresponding to each U_p-eigenvalue, and hence the entire q-expansion of each eigenform (other than the coefficients at primes dividing N). However, computing the trace of T_m grows very rapidly with m, and consequently this is not a very practical algorithm.

Moreover, there is no reason why k_∞ should be an integer; it is sufficient to take k_∞ to be an arbitrary continuous homomorphism $\mathbb{Z}_p^* \to \mathbb{C}_p^*$. The space $\mathcal{W}$ of such homomorphisms is naturally a rigid space, isomorphic to a disjoint union of $p - 1$ open discs of radius 1, where the homomorphism κ corresponds to $w = \kappa(1 + p) - 1$. (For $p = 2$ one takes instead 2 open discs of radius 1, with $w = \kappa(1 + 2^2) - 1$.) Coleman has shown how to construct a space of overconvergent forms of weight κ for arbitrary $\kappa \in \mathcal{W}$, and the trace formula method applies to these p-adic weights just as before. Indeed, one can easily obtain a formula for $\mathrm{Tr}(U_p|S_\kappa)$, for κ in a fixed component of weight space, as a power series in W; this is useful for studying local geometry of eigenvarieties.

References

[Eme] Matthew Emerton, *p-adic geometry of modular curves*, Online notes.

[Kat73] Nicholas M. Katz, *p-adic properties of modular schemes and modular forms*, Modular functions of one variable III (Antwerp, 1972), Lecture Notes in Mathematics, vol. 350, Springer, 1973, pp. 69–190. MR 0447119

[Kol61] Oddmund Kolberg, *Congruences for the coefficients of the modular invariant $j(\tau)$ modulo powers of 2*, Årbok Univ. Bergen Mat.-Natur. Ser. (1961), 1–9.

[Köp74] Ursula Köpf, *Über eigentliche Familien algebraischer Varietäten über affinoiden Räumen*, Schr. Math. Inst. Univ. Münster (2nd ser.) **7** (1974), iv+72. MR 0422671

[Loe07] David Loeffler, *Spectral expansions of overconvergent modular functions*, Int. Math. Res. Notices **2007** (2007), no. 050. MR 2353090

[Ser73] Jean-Pierre Serre, *Formes modulaires et fonctions zeta p-adiques*, Modular functions of one variable III (Antwerp, 1972), Lecture Notes in Mathematics, vol. 350, Springer, 1973, pp. 191–268. MR 0404146