Pure and Applied Mathematics Quarterly
Volume 1, Number 1, 205—221, 2005

MODULAR FORMS FOR NONCONGRUENCE SUBGROUPS

Wen-Ching Winnie Li, Ling Long and Zifeng Yang

1 Introduction

The study of modular forms for congruence subgroups of $SL_2(\mathbb{Z})$ has been one of the central topics in number theory for over one century. It has broad applications and impact to many branches of mathematics. Langlands' program is a vast generalization of this subject from representation-theoretic point of view. The most recent highlight is the proof of the Taniyama-Shimura-Weil modularity conjecture by Wiles [Wi], Taylor-Wiles [TW] , and Breuil-Conrad-Diamond-Taylor [BCDT], which leads to the establishment of Fermat's last theorem.

The story for noncongruence subgroups of $SL_2(\mathbb{Z})$ is totally different. Despite of the fact that there are far more noncongruence subgroups than congruence subgroups, as pointed out by Jones in [J1, J2], and the fact that any smooth irreducible projective curve defined over a number field is birationally equivalent to a modular curve, mostly from a noncongruence subgroup, as proved by Belyi [Bel], modular forms for noncongruence subgroups have not attracted wide attention. The main contributors to this topic are Atkin, Swinnerton-Dyer, and Scholl. Atkin and Swinnerton-Dyer [ASD] pioneered the research in this area; they laid foundations and made a remarkable observation on the congruence property of Fourier coefficients of cusp forms for noncongruence subgroups, which we call Atkin-Swinnerton-Dyer congruences. Scholl [Sc1] attaches to the space of cusp forms of a given weight for a noncongruence subgroup a compatible family of l-adic Galois representations and proves that the Fourier coefficients of all cusp forms in the space satisfy certain congruence relations arising from the characteristic polynomials of the Frobenius elements. In case the space is one-dimensional,

Received November 16, 2004.
The research of the first author was supported in part by an NSA grant MDA904-03-1-0069.

Scholl's result implies the Atkin-Swinnerton-Dyer congruences. Moreover, Scholl proves in [Sc4, Sc5] that in a few one-dimensional cases, his l-adic representations are modular so that the Atkin-Swinnerton-Dyer congruences give congruence relations between noncongruence cusp forms and congruence cusp forms.

In [LLY] we investigate an example of noncongruence subgroup whose weight 3 space of cusp forms is two-dimensional. We show that the l-adic Galois representation attached to this space by Scholl actually decomposes over a quadratic extension into the sum of two representations associated to two cuspidal newforms of congruence subgroups, the Atkin-Swinnerton-Dyer congruences hold for this space, and it asserts congruence relations between noncongruence cusp forms and congruence cusp forms.

This is a survey article aimed at readers interested in modular forms. In §2 we recall the main arithmetic features of cuspidal newforms for congruence subgroups, which are to be contrasted in §3 with cusp forms for noncongruence subgroups. The Atkin-Swinnerton-Dyer congruences and Scholl's Galois representations are explained in detail there. The main results of [LLY] are reviewed in §4 with a sketch of proof given in §5. Our proof of modularity of the Galois representation uses the Faltings-Serre criterion [Ser] since we want to show it isomorphic to a known representation. It would be very interesting to obtain the modularity conclusion by an alternative method, using the criterion of Skinner-Wiles given in [SW].

This paper grows out of the invited talk given by the first author in the International Conference in Memory of Armand Borel, July 26-30, 2004, held at the Center of Mathematical Sciences at Zhejiang University. She would like to thank the Center for its hospitality.

2 Modular forms for congruence subgroups

The modular forms for congruence subgroups of $SL_2(\mathbb{Z})$ are well-understood. The theory of newforms describes the structure of such forms as well as the arithmetic properties of the Fourier coefficients of newforms. Specifically, let $g = \sum_{n\geq 1} a_n(g)q^n$, where $q = e^{2\pi i\tau}$, be a cuspidal newform of weight $k \geq 2$ level N and character χ, normalized with the leading Fourier coefficient $a_1(g) = 1$. We summarize below the main properties of the Fourier coefficients of g. The reader is referred to [AL] and [Li] for more details on the theory of newforms.

- Viewed as an arithmetic function on positive integers, it is multiplicative,

that is,

$$a_{mn}(g) = a_m(g)a_n(g)$$

whenever m and n are coprime.

So it remains to explain how the Fourier coefficients behave with respect to varying primes powers.

- g is an eigenfunction of the Hecke operators T_p with eigenvalue $a_p(g)$ for all primes $p \nmid N$. By a theorem of Hecke [He], this translates as

$$a_{np}(g) - a_p(g)a_n(g) + \chi(p)p^{k-1}a_{n/p}(g) = 0 \tag{1}$$

for all primes p not dividing N and for all $n \geq 1$.

For primes $p|N$, we have

$$a_{np}(g) = a_n(g)a_p(g)$$

for all $n \geq 1$.

- The Fourier coefficients of a newform are algebraic integers. Further, for a congruence subgroup and a given weight, there is a basis of cusp forms with integral Fourier coefficients. Hence the Fourier coefficients of a cusp form, if all algebraic, have bounded denominators.

- Thanks to the work of Eichler, Shimura, and Deligne [Ei,Sh1,del1], there exists a compatible family of irreducible λ-adic representations $\rho_{\lambda,g}$ of the Galois group $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$, unramified outside lN, where λ divides l, such that

$$\begin{aligned}\mathrm{Tr}(\rho_{\lambda,g}(\mathrm{Frob}_p)) &= a_p(g),\\ \det(\rho_{\lambda,g}(\mathrm{Frob}_p)) &= \chi(p)p^{k-1},\end{aligned}$$

for all primes p not dividing lN. Therefore the characteristic polynomial $H_p(T) = T^2 - A_1(p)T + A_2(p)$ of $\rho_{\lambda,g}(\mathrm{Frob}_p)$ is independent of the λ's not dividing p, and the Fourier coefficients of g satisfy the relation

$$a_{np}(g) - A_1(p)a_n(g) + A_2(p)a_{n/p}(g) = 0 \tag{2}$$

for all $n \geq 1$ and all primes p not dividing N.

- A consequence of the above is the following estimate of the Fourier coefficients $a_p(g)$ for primes p not dividing the level N. They satisfy the Ramanujan-Petersson conjecture

$$|a_p(g)| \leq 2p^{(k-1)/2}.$$

These are the key properties to be compared with cusp forms for noncongruence subgroups.

There is a parallel theory of newforms on Eisenstein series of congruence subgroups; the associated λ-adic representations are reducible, and the Ramanujan-Petersson conjecture does not hold.

3 Modular forms for noncongruence subgroups

The study of modular forms for noncongruence subgroups was initiated by Atkin and Swinnerton-Dyer [ASD], and continued in a sequence of papers [Sc1, Sc3, Sc4, Sc5, Sc6] by Scholl. Compared with what we know for forms for congruence subgroups, the overall knowledge on the arithmetic properties of forms for noncongruence subgroups, however, is far from satisfactory. Let Γ be a noncongruence subgroup of $SL_2(\mathbb{Z})$ with finite index. Denote by $S_k(\Gamma)$ the space of cusp forms of weight $k \geq 2$ for Γ. It is finite-dimensional, denoted by d. Let μ denote the width of the cusp at ∞. A cusp form for Γ has an expansion in powers of $q^{1/\mu}$, just like forms for congruence subgroups. The similarity more or less stops here. We proceed to discuss the dissimilarities.

Scholl in [Sc2] studied Eisenstein series for noncongruence subgroups. Unlike the case for congruence subgroups, he discovered that the arithmetic properties for noncongruence Eisenstein series are quite unclear. In this paper we shall confine ourselves to cusp forms for noncongruence subgroups. We begin with the integrality property. Through the work of Atkin and Swinnerton-Dyer [ASD] and Scholl [Sc1] it is known that the modular curve X_Γ, which is the quotient of the Poincaré upper half-plane by Γ compactified by joining the cusps of Γ, has an algebraic model defined over a number field K. Further, they show that there exists a subfield L of K, an element $\kappa \in K$ with $\kappa^\mu \in L$, and a positive integer M such that κ^μ is integral outside M and $S_k(\Gamma)$ has a basis consisting of M-integral forms. Here a modular form f is said to be M-integral if in its Fourier expansion

$$f(\tau) = \sum_{n \geq 1} a_n(f) q^{n/\mu},$$

the Fourier coefficients $a_n(f)$ can be written as $\kappa^n c_n(f)$ with $c_n(f)$ lying in L, integral everywhere except possibly at the places dividing M.

The paper by Scholl [Sc6] and Serre's letter to Thompson in [Th] as well as its generalization by Berger [Br1] indicate that the obvious extension of Hecke operators to noncongruence subgroups would not reveal much information on forms which genuinely live on noncongruence subgroups. Hence there are no parallel Hecke theory on forms for noncongruence subgroups giving similar arithmetic information on Fourier coefficients. On the other hand, based on their handful numerical data, Atkin and Swinnerton-Dyer [ASD] made an amazing observation of congruence relations for certain cusp forms for noncongruence subgroups, from which and our own numerical data, we formulate the following conjecture.

Conjecture 3.1 (Atkin-Swinnerton-Dyer congruences)**.** Suppose that the modular curve X_Γ has a model over $\mathbb{Q}$ in the sense of [§5][Sc1]. There exist a positive integer M and a basis of $S_k(\Gamma)$ consisting of M-integral forms f_j, $1 \le j \le d$, such that for each prime p not dividing M, there exists a nonsingular $d \times d$ matrix $(\lambda_{i,j})$ whose entries are in a finite extension of $\mathbb{Q}_p$, algebraic integers $A_p(j)$, $1 \le j \le d$, with $|\sigma(A_p(j))| \le 2p^{(k-1)/2}$ for all embeddings σ, and characters χ_j unramified outside M so that for each j the Fourier coefficients of $h_j := \sum_i \lambda_{i,j} f_i$ satisfy the congruence relation

$$\mathrm{ord}_p(a_{np}(h_j) - A_p(j)a_n(h_j) + \chi_j(p)p^{k-1}a_{n/p}(h_j)) \ge (k-1)(1+\mathrm{ord}_p n) \quad (3)$$

for all $n \ge 1$; or equivalently, for all $n \ge 1$,

$$(a_{np}(h_j) - A_p(j)a_n(h_j) + \chi_j(p)p^{k-1}a_{n/p}(h_j))/(np)^{k-1}$$

is integral at all places dividing p.

In other words, the recursive relation (1) on Fourier coefficients of modular forms for congruence subgroups is replaced by the congruence relation (3) for forms of noncongruence subgroups. The meaning of $A_p(j)$'s is mysterious. Note that they satisfy the Ramanujan-Petersson bound. The examples in [ASD] suggest that they verify the Sato-Tate conjecture.

Regarding the Atkin-Swinnerton-Dyer conjecture, Scholl in [Sc1] proved the following result.

Theorem 3.1 (Scholl)**.** *Suppose that X_Γ has a model over $\mathbb{Q}$ as before. Attached to $S_k(\Gamma)$ is a compatible family of $2d$-dimensional l-adic representations ρ_l of the Galois group $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ unramified outside lM such that for primes $p > k+1$ not dividing Ml, the following hold.*

(i) The characteristic polynomial

$$H_p(T) = \sum_{0 \le r \le 2d} B_r(p) T^{2d-r} \tag{4}$$

of $\rho_l(\mathrm{Frob}_p)$ *lies in* $\mathbb{Z}[T]$ *and is independent of* l*, and its roots are algebraic integers with absolute value* $p^{(k-1)/2}$*;*

(ii) For any M*-integral form* f *in* $S_k(\Gamma)$*, its Fourier coefficients* $a_n(f)$*,* $n \ge 1$*, satisfy the congruence relation*

$$\begin{aligned} &\mathrm{ord}_p(a_{np^d}(f) + B_1(p)a_{np^{d-1}}(f) + \cdots + B_{2d-1}(p)a_{n/p^{d-1}}(f) + B_{2d}(p)a_{n/p^d}(f)) \\ &\ge (k-1)(1 + \mathrm{ord}_p n) \end{aligned} \tag{5}$$

for $n \ge 1$*.*

Scholl's theorem establishes the Atkin-Swinnerton-Dyer congruences if $S_k(\Gamma)$ has dimension 1. If the Atkin-Swinnerton-Dyer congruences were established in general, then

$$H_p(T) = \prod_{1 \le j \le d} (T^2 - A_p(j)T + \chi_j(p)p^{k-1}).$$

Therefore Scholl's congruence relation (5) may be regarded as a collective replacement for forms of noncongruence subgroup of the equality (2) for newforms. Likewise, the representations ρ_l are a collective replacement of λ-adic representations attached to newforms for congruence subgroups.

4 Atkin-Swinnerton-Dyer congruence relations

If Scholl's degree $2d$ representation ρ_l were decomposable as a direct sum of d degree 2 representations of the Galois group $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$, and each of these were proven to be modular, that is, arising from a newform of weight k as predicted by the Langlands philosophy, then, combined with the Atkin-Swinnerton-Dyer conjecture, one would obtain congruence relations between cusp forms for congruence and noncongruence groups. We explain this in more detail.

Let $f = \sum_{n \ge 1} a_n(f) q^{n/\mu}$ be an M-integral cusp form in $S_k(\Gamma)$, and let $g = \sum_{n \ge 1} b_n(g) q^n$ be a normalized newform of weight k level N and character χ.

Definition 4.1. The two forms f and g above are said to satisfy the Atkin-Swinnerton-Dyer [ASD] congruence relations if, for all primes p not dividing MN and for all $n \geq 1$,

$$(a_{np}(f) - b_p(g)a_n(f) + \chi(p)p^{k-1}a_{n/p}(f))/(np)^{k-1} \tag{6}$$

is integral at all places dividing p.

Not much is known about ASD congruence relations. In [ASD] the noncongruence subgroup $\Gamma_{7,1,1}$ was studied. This is an index 9 subgroup of $SL_2(\mathbb{Z})$ with three cusps of cusp width 7, 1, 1, respectively. Its space of cusp forms of weight 4 is one-dimensional, containing a nonzero 14-integral form f. Scholl proved in [Sc4] that the l-adic representation he attached to the space $S_4(\Gamma_{7,1,1})$ is the l-adic representation associated to a normalized newform g of weight 4 level 14 and trivial character. Therefore, by Scholl's theorem, f and g satisfy the ASD congruence relations. In the unpublished paper [Sc5], Scholl obtained a similar result for $S_4(\Gamma_{4,3})$ and $S_4(\Gamma_{5,2})$; both spaces are also 1-dimensional.

Denote by $\Gamma^1(5)$ the congruence subgroup

$$\Gamma^1(5) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ * & 1 \end{pmatrix} \mod 5 \right\}.$$

It has four cusps $\infty, -2, 0, -5/2$ with cusp widths 5, 5, 1, 1, respectively. In our recent joint paper [LLY], we study the noncongruence subgroup Γ, an index three normal subgroup of $\Gamma^1(5)$ in which the cusps ∞ and -2 of $\Gamma^1(5)$ are totally ramified, while the other two cusps of $\Gamma^1(5)$ split completely. The space $S_3(\Gamma)$ has dimension two. The following detailed description about $S_3(\Gamma)$ shows in particular that ASD congruence relations hold for this space. It is also proved in [LLY] that the same hold for the one-dimensional weight 3 cusp forms for the subgroup G of $SL_2(\mathbb{Z})$ such that there is an elliptic modular $K3$ surface defined over $\mathbb{Q}$ fibered over the curve X_G.

Theorem 4.1 (Main Theorem)**.** *Let Γ be the index 3 noncongruence subgroup of $\Gamma^1(5)$ such that the two cusps ∞ and -2 of $\Gamma^1(5)$ are totally ramified and the other two cusps of $\Gamma^1(5)$ split completely.*

(1) Then X_Γ has a model over $\mathbb{Q}$, $\kappa = 1$, and the space $S_3(\Gamma)$ is 2-dimensional

with a basis consisting of 3*-integral forms*

$$\begin{aligned} f_+(\tau) &= q^{1/15} + iq^{2/15} - \frac{11}{3}q^{4/15} - i\frac{16}{3}q^{5/15} - \frac{4}{9}q^{7/15} + i\frac{71}{9}q^{8/15} \\ &\quad + \frac{932}{81}q^{10/15} + O(q^{11/15}), \\ f_-(\tau) &= q^{1/15} - iq^{2/15} - \frac{11}{3}q^{4/15} + i\frac{16}{3}q^{5/15} - \frac{4}{9}q^{7/15} - i\frac{71}{9}q^{8/15} \\ &\quad + \frac{932}{81}q^{10/15} + O(q^{11/15}). \end{aligned}$$

(2) *The* 4*-dimensional* l*-adic representation* ρ_l *of* $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ *associated to* $S_3(\Gamma)$ *constructed by Scholl is modular. More precisely, there are two cuspidal newforms of weight 3 level 27 and character* χ_{-3} *given by*

$$\begin{aligned} g_+(\tau) &= q - 3iq^2 - 5q^4 + 3iq^5 + 5q^7 + 3iq^8 + 9q^{10} + 15iq^{11} - 10q^{13} - 15iq^{14} \\ &\quad -11q^{16} - 18iq^{17} - 16q^{19} - 15iq^{20} + 45q^{22} + 12iq^{23} + O(q^{24}) \\ g_-(\tau) &= q + 3iq^2 - 5q^4 - 3iq^5 + 5q^7 - 3iq^8 + 9q^{10} - 15iq^{11} - 10q^{13} + 15iq^{14} \\ &\quad -11q^{16} + 18iq^{17} - 16q^{19} + 15iq^{20} + 45q^{22} - 12iq^{23} + O(q^{24}) \end{aligned}$$

such that over the extension of $\mathbb{Q}_l$ *by joining* $\sqrt{-1}$*,* ρ_l *decomposes into the direct sum of the two* λ*-adic representations attached to* g_+ *and* g_-*, where* λ *is a place of* $\mathbb{Q}(i)$ *dividing* l*.*

(3) f_+ *and* g_+ *(resp.* f_- *and* g_-*) satisfy the Atkin-Swinnerton-Dyer congruence relations.*

Here χ_{-3} is the quadratic character attached to the field $\mathbb{Q}(\sqrt{-3})$.

The statement (3) not only establishes the Atkin-Swinnerton-Dyer conjecture by finding a basis independent of p, but also explains the mysterious $A_p(j)$ in the conjecture by showing that, in this case, they come from newforms for a congruence subgroup!

Since the group Γ does not contain the minus identity matrix in $SL_2(\mathbb{Z})$, there is an elliptic surface $\mathcal{E}$ fibred over the modular curve X_Γ. The statement (2) then implies that the L-function attached to the surface $\mathcal{E}$ is a quotient of two automorphic L-functions.

The fact that ρ_l decomposes over $\mathbb{Q}_l(\sqrt{-1})$ as stated in (2) may be a special phenomenon. Scholl proved in [Sc7] that the l-adic representations attached to $S_4(\Gamma_{7,1,1}), S_4(\Gamma_{4,3})$ and $S_4(\Gamma_{5,2})$ have images which are symplectic groups. This is partly due to the presence of an alternating pairing in case of even weight. The pairing is symmetric for odd weight case.

5 A sketch of a proof of the main theorem

It follows easily from the dimension formula in Shimura [Sh2] that the space $S_3(\Gamma)$ has dimension two. Denote by E_1 (resp. E_2) the Eisenstein series of weight 3 for $\Gamma^1(5)$ which vanishes at all cusps of $\Gamma^1(5)$ except with value 1 at the cusp ∞ (resp. -2). Then they have no other zeros on the modular curve $X^1(5)$. Since the curve X_Γ is a 3-fold cover of $X^1(5)$ totally ramified at the cusps ∞ and -2 and unramified elsewhere, we may take cubic roots of $E_1^2E_2$ and $E_1E_2^2$, denoted by f_1 and f_2, respectively, to get cusp forms in $S_3(\Gamma)$. The forms $f_\pm$ in the theorem are linear combinations of f_1 and f_2, namely, $f_\pm = f_1 \pm if_2$. Observe that $f_\pm$ have algebraic Fourier coefficients, which are integral outside 3, and 3 appears in the denominators of these Fourier coefficients with exponents apparently going to infinity. This unbounded denominator phenomenon shows that Γ is a noncongruence subgroup.

The modular curve $X^1(5)$ has genus zero, and so does X_Γ. We may choose

$$t = \frac{f_1}{f_2} = \sqrt[3]{\frac{E_1}{E_2}}$$

as a generator of the function field of X_Γ. In other words, t identifies the modular curve X_Γ with the projective line $\mathbb{P}^1$. There is an elliptic surface $\mathcal{E}_\Gamma$ fibered over X_Γ, defined by the explicit equation

$$y^2 + (1-t^3)xy - t^3y = x^3 - t^3x^2\,. \tag{7}$$

The matrix $A = \begin{pmatrix} -2 & -5 \\ 1 & 2 \end{pmatrix}$ normalizes the group $\Gamma^1(5)$ and Γ; by mapping t to $-1/t$ it induces an involution over $\mathbb{Q}$ on the modular curve X_Γ and an order 4 action on the elliptic surface $\mathcal{E}$.

The bulk of the proof is devoted to showing the modularity of ρ_l attached to $S_3(\Gamma)$, that is, over $\mathbb{Q}_l(i)$ it decomposes into the sum of the λ-adic representations attached to g_+ and g_-. We first describe Scholl's representation ρ_l. Choose an integer $N \geq 3$ such that

$$\pm\Gamma\ \Gamma(N) = SL_2(\mathbb{Z}).$$

Write $X(N)^0$ for $X(N)$ - {cusps}, and $V(N)$ for the normalization of the fiber product $X_\Gamma \times_{X(1)} X(N)$, the curve corresponding to the intersection of Γ with $\Gamma(N)$. Denote by $V(N)^0$ the normalization of the fiber product $X_\Gamma \times_{X(1)} X(N)^0$. Let $G(N) = SL(\mu_N \times \mathbb{Z}/N)$. Since a universal elliptic curve is known to exist over a principal congruence modular curve $X(N)$, but not over X_Γ, the idea is to use the one over $X(N)$ to construct a sheaf over the intersection $V(N)$, then to get rid of the dependence of $\Gamma(N)$ by taking $G(N)$-invariants. To carry

this out technically, let $f^{\text{univ}} : E^{\text{univ}} \to X(N)^0$ be the restriction to $X(N)^0$ of the universal elliptic curve of $X(N)$. Starting with the constant sheaf $\mathbb{Q}_l$ on the universal elliptic curve E^{univ}, pushing it down using f_*^{univ}, and applying the derived functor, we obtain a sheaf

$$\mathcal{F}_l^{\text{univ}} = R^1 f_*^{\text{univ}} \mathbb{Q}_l$$

on $X(N)^0$. The image of this under the pull back of the natural projection $\pi_0' : V(N)^0 \to X(N)^0$ followed by the push forward of the injection map $i_N : V(N)^0 \to V(N)$ is a sheaf $(i_N)_* \pi_0'^* \mathcal{F}_l^{\text{univ}}$ on $V(N)$. Scholl's representation ρ_l is the action of $\text{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ on the $G(N)$-invariant $\mathbb{Q}_l$-space which is the first étale cohomology group of the $\bar{\mathbb{Q}}$-rational points on $V(N)$:

$$H^1(V(N) \otimes \bar{\mathbb{Q}}, (i_N)_* \pi_0'^* \mathcal{F}_l^{\text{univ}})^{G(N)}.$$

We take advantage of the elliptic surface

$$h : \mathcal{E}_\Gamma \to X_\Gamma$$

to define a less complicated l-adic representation ρ_l^* of $\text{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ as follows. Write X_Γ^0 for X_Γ - {cusps}, and denote by $i : X_\Gamma^0 \to X_\Gamma$ the inclusion map. By abuse of notation, the restriction map $\mathcal{E}_\Gamma \to X_\Gamma^0$ is also denoted by h. We get a sheaf

$$\mathcal{F}_l = R^1 h_* \mathbb{Q}_l$$

on X_Γ^0 obtained by applying the derived functor R^1 to the push forward by h of the constant sheaf $\mathbb{Q}_l$ on $\mathcal{E}_\Gamma$. The action of the Galois group on

$$W_l = H^1(X_\Gamma \otimes \overline{\mathbb{Q}}, i_* \mathcal{F}_l)$$

is the representation ρ_l^*. It is not hard to show that ρ_l^* and ρ_l are isomorphic up to a twist by a character ϕ_l of order at most 2. For the rest of the proof, take $l = 2$.

The symmetry with respect to A allows us to regard ρ_2^* as a degree 2 representation over $K = \mathbb{Q}_2(A)$. We compute the characteristic polynomial $H_p(T)$ (resp. $H_p'(T)$) of $\rho_2^*(\text{Frob}_p)$ over $\mathbb{Q}_2$ (resp. K) using the Lefschetz formula. This amounts to computing $\text{Tr}(\rho_2^*(\text{Frob}_q))$ for $q = p, p^2$. Aided by a computer, the computation is performed for $p = 5, ..., 23$, which enables us to obtain $H_p(T)$ completely, and determine $H_p'(T)$ except when $p \equiv 2 \pmod 3$, in which case its linear term is determined up to sign. The result is tabulated below.

p	$H_p(T)$	$H'_p(T)$
5	$T^4 - 41\,T^2 + 625$	$T^2 \pm 3A\,T - 25$
7	$T^4 - 10\,T^3 + 123\,T^2 - 490\,T + 7^4$	$T^2 - 5\,T + 7^2$
11	$T^4 - 17\,T^2 + 11^4$	$T^2 \pm 15A\,T - 11^2$
13	$T^4 + 20\,T^3 + 438\,T^2 + 20 \cdot 13^2\,T + 13^4$	$T^2 + 10\,T + 13^2$
17	$T^4 - 254\,T^2 + 17^4$	$T^2 \pm 18A\,T - 17^2$
19	$T^4 + 32\,T^3 + 978\,T^2 + 32 \cdot 19^2\,T + 19^4$	$T^2 + 16\,T + 19^2$
23	$T^4 - 914\,T^2 + 23^4$	$T^2 \pm 12A\,T - 23^2$

The information above also allows us to conclude that the two-dimensional representation ρ_2^* over K has its determinant equal to $\chi_{-3}\varepsilon_2^2$, where ε_2 denotes the 2-adic cyclotomic character.

Now let $\tilde{\rho}_2$ denote the 4-dimensional 2-adic representation of the Galois group of $\mathbb{Q}$ attached to $g_+ \oplus g_-$. The Fricke involution H_{27} exchanges $g_\pm$ up to sign and is $\mathbb{Q}$-rational. So $\tilde{\rho}_2$ may be regarded as a 2-dimensional representation over $\mathbb{Q}_2(i)$, which is equal to $\mathbb{Q}(i)$ completed at $1+i$, denoted by $\mathbb{Q}(i)_{1+i}$.

Fix an isomorphism from K to $\mathbb{Q}_2(i)$ such that the characteristic polynomial of $\rho_2^*(\mathrm{Frob}_5)$ agrees with that of $\tilde{\rho}_2(\mathrm{Frob}_5)$. Regard ρ_2^* as a 2-dimensional representation over $\mathbb{Q}(i)_{1+i}$. The goal is to show that ρ_2^* and $\tilde{\rho}_2$ are isomorphic 2-dimensional representations over $\mathbb{Q}_2(i)$. We present the main steps below.

(i) $\det \rho_2^* = \det \tilde{\rho}_2 = \chi_{-3}\varepsilon_2^2$.

(ii) Information on H'_p for $p = 5, 7, 13$ implies that, modulo $1+i$, ρ_2^* and $\tilde{\rho}_2$ are isomorphic representations, with image equal to $\mathrm{GL}_2(\mathbb{F}_2)$.

We can then invoke Faltings-Serre criterion [Ser], which says that under (i) and (ii), if ρ_2^* and $\tilde{\rho}_2$ are not isomorphic, then there exists a "deviation" group, constructed from the two representations, which measures the difference of the two representations. In the case present, the "deviation" group arises from a Galois extension of $\mathbb{Q}$ unramified outside 2 and 3, and is isomorphic to $S_4 \times \{\pm 1\}$, or S_4, or $S_3 \times \{\pm 1\}$, such that the two representations have the same traces on Frobenius elements with order at most three, and different traces on all Frobenius elements of order at least 4. Hence it suffices to show

(iii) For each Galois extension of $\mathbb{Q}$ unramified outside 2 and 3, and with Galois group isomorphic to $S_4 \times \{\pm 1\}$, or S_4, or $S_3 \times \{\pm 1\}$, there is a Frobenius element of order at least 4 at which both representations have the same trace.

The statement (iii) is proved as follows. From analyzing the representations mod $1+i$ we know that the fixed field of the possible deviation group contains

the field

$$K_6 = \mathrm{Split}(x^3 + 3x - 2) \supset \mathbb{Q}(\sqrt{-6}).$$

Start by finding all fields M with $\mathrm{Gal}(M/\mathbb{Q})$ equal to S_4, which contain K_6 and are unramified outside 2 and 3. There are three such fields.

defining equation	discriminant	p with order 4 Frob_p
$x^4 - 4x - 3 = 0$	$-2^9 \cdot 3^3$	13, 17, 19, 23
$x^4 - 8x + 6 = 0$	$-2^1 3 \cdot 3^3$	13, 17
$x^4 - 12x^2 - 16x + 12 = 0$	$-2^1 3 \cdot 3^3$	19, 23

Since $H'_p(T)$ agrees with the characteristic polynomial of $\tilde{\rho}_2^*(\mathrm{Frob}_p)$ for $p = 13, 19$, we rule out deviation group isomorphic to $S_4 \times \{\pm 1\}$ and S_4. To rule out the last case $S_3 \times \{\pm 1\}$, we consider Frob_p for $p = 5, 7$. On the one hand, such Frob_p has order 3 in S_3 and it has order 2 in a quadratic extension unramified outside 2 and 3. Hence it has order 6 in $S_3 \times \{\pm 1\}$. On the other hand, both representations have the same trace on these two Frob_p by construction.

Finally, we apply Scholl's congruence relation and use the explicit characteristic polynomial $H_p(T)$ at $p = 7, 13$ to show ϕ_2 is either trivial or χ_{-3}. Since $\tilde{\rho}_2$ is invariant under the twist by χ_{-3}, we have

$$\rho_2 \cong \rho_2^* \cong \tilde{\rho}_2.$$

This proves the second statement.

It remains to prove the ASD congruence relations. This is obtained by comparing the p-adic theory from crystalline cohomology and the dyadic theory as follows. Fix a prime $p \neq 2, 3$.

Recall that the operator A on $S_3(\Gamma)$ satisfies

$$A(f_1) = f_2, \qquad A(f_2) = -f_1.$$

There is another map ξ on $S_3(\Gamma)$ arising from the map sending t to $\omega^2 t$, where $\omega = e^{2\pi i/3}$, on the elliptic surface $\mathcal{E}$:

$$\xi(f_1) = \omega f_1, \qquad \xi(f_2) = \omega^2 f_2.$$

The eigenvalues of A are i and $-i$, and those of ξ are ω and ω^2. We consider three spaces, each of which will be decomposed as a direct sum of eigenspaces with eigenvalues appearing as subindices.

(a) Space V for p-adic theory

$$V = S_3(\Gamma, \mathbb{Z}_p) \oplus S_3(\Gamma, \mathbb{Z}_p)^\vee = V_i \oplus V_{-i} = V_\omega \oplus V_{\omega^2},$$

where

$$\begin{aligned} V_i &= \langle f_1 - if_2 = f_-, f_1^\vee - if_2^\vee \rangle, \\ V_{-i} &= \langle f_1 + if_2 = f_+, f_1^\vee + if_2^\vee \rangle, \\ V_\omega &= \langle f_1, f_2^\vee \rangle, \quad V_{\omega^2} = \langle f_2, f_1^\vee \rangle. \end{aligned}$$

Denote the action of the Frobenius by F.

(b) Space W of the representation ρ_2

We have similar decomposition of W as a direct sum of eigenspaces of A and of ξ, respectively:

$$W = W_i \oplus W_{-i} = W_\omega \oplus W_{\omega^2},$$

Denote by F_p the action of the Frobenius.

(c) Space $\tilde{W}$ of $\tilde{\rho}_2$

It is the 4-dimensional 2-adic space attached to g_+ and g_- on which the operator

$$H_{27} = \begin{pmatrix} 0 & -1 \\ 27 & 0 \end{pmatrix}$$

plays the role of A. Similar decomposition holds on $\tilde{W}$:

$$\tilde{W} = \tilde{W}_i \oplus \tilde{W}_{-i}.$$

Note that $g_- \in \tilde{W}_i$ and $g_+ \in \tilde{W}_{-i}$. Denote by Frob_p the action of the Frobenius automorphism.

Write $char(U, T)$ for the characteristic polynomial of the operator T on the space U. The isomorphism $\rho_2 \cong \tilde{\rho}_2$ implies

$$char(W_{\pm i}, F_p) = char(\tilde{W}_{\pm i}, \mathrm{Frob}_p).$$

Further, Scholl has shown in [Sc1] that

$$char(V, F) = char(W, F_p).$$

We prove a refinement of this relation from which the stated Atkin-Swinnerton-Dyer congruence relations follow.

Theorem 5.1. *There hold*

$$char(V_i, F) = char(W_i, F_p) = char(\tilde{W}_i, \mathrm{Frob}_p),$$
$$char(V_{-i}, F) = char(W_{-i}, F_p) = char(\tilde{W}_{-i}, \mathrm{Frob}_p).$$

To prove this, first observe the commutativity relations among the operators : $AF = FA, AF_p = F_pA, A\xi = \xi^2 A$. When $p \equiv 1 \mod 3$, the operator A maps V_ω isomorphically onto V_{ω^2}, while F preserves V_ω and V_{ω^2}. This implies $char(V, F) = char(V_\omega, F)^2$. Further, V_i and V_{-i} are also F-invariant. By analyzing the eigenvalues of F on V_ω and V_i, we conclude that

$$char(V_i, F) = char(V_{-i}, F) = char(V_\omega, F) = char(V_{\omega^2}, F).$$

On the dyadic side, we use the space $\tilde{W}$. Noting that g_+ is the twist of g_- by χ_{-3} and $\chi_{-3}(p) = 1$ for primes $p \equiv 1 \pmod 3$, we get $char(\tilde{W}_i, \mathrm{Frob}_p) = char(\tilde{W}_{-i}, \mathrm{Frob}_p)$, and consequently

$$char(\widetilde{W}_i, \mathrm{Frob}_p) = char(W_i, F_p) = char(V_i, F),$$

as desired.

For $p \equiv 2 \mod 3$, we have

$$char(V_i, FA) = char(V_{-i}, FA) = char(V_\omega, FA) = char(V_{\omega^2}, FA)$$

and on dyadic side, we have

$$char(W_i, F_pA) = char(\widetilde{W}_i, \mathrm{Frob}_p H_{27})$$
$$= char(\widetilde{W}_{-i}, \mathrm{Frob}_p H_{27}) = char(W_{-i}, F_pA).$$

We then use $char(V, FA) = char(W, F_pA)$ proved by Scholl to conclude

$$char(V_i, FA) = char(W_i, F_pA).$$

Since the action of A on both spaces is multiplication by i, this implies

$$char(V_i, F) = char(W_i, F_p).$$

References

[AL] A. O. L. Atkin and J. Lehner, Hecke operators on $\Gamma_0(m)$, *Math. Ann.* **185** (1970), 134-160.

[ASD] A. O. L. Atkin and H. P. F. Swinnerton-Dyer, Modular forms on noncongruence subgroups, *Combinatorics* (Proc. Sympos. Pure Math., Vol. XIX, Univ. California, Los Angeles, Calif., 1968), Amer. Math. Soc., Providence, R.I., 1971, pp. 1–25.

[Bel] G. V. Belyĭ, Galois extensions of a maximal cyclotomic field, *Izv. Akad. Nauk SSSR Ser. Mat.* **43** (1979), no. 2, 267–276, 479.

[Br1] G. Berger, Hecke operators on noncongruence subgroups, *C. R. Acad. Sci. Paris Sér. I Math.*, **319** (1994), no. 9, 915-919.

[BCDT] C. Breuil, B. Conrad, F. Diamond, and R. Taylor, On the modularity of elliptic curves over $\mathbb{Q}$: wild 3-adic exercises, *J. Amer. Math. Soc* **14** (2001), no. 4, 843-939.

[De] P. Deligne, Formes modulaires et représentations de gl(2), *Modular functions of one variable*, II (Proc. Internat. Summer School, Univ. Antwerp, Antwerp, 1972), Springer, Berlin, 1973, pp. 55–105. Lecture Notes in Math., Vol. 349.

[Ei] M. Eichler, Eine Verallgemeinerung der Abelschen Integrale, *Math. Z.* **67** (1957), 267–298.

[J1] Gareth A. Jones, Triangular maps and noncongruence subgroups of the modular group, *Bull. London Math. Soc.* **11** (1979), no. 2, 117–123.

[J2] Gareth A. Jones, Congruence and noncongruence subgroups of the modular group: a survey, Proceedings of groups—St. Andrews 1985 (Cambridge), *London Math. Soc. Lecture Note Ser.*, vol. **121**, Cambridge Univ. Press, 1986, pp. 223–234.

[He] E. Hecke, *Mathematische Werke*. Göttingen: Vandenhoeck und Ruprecht. (1959).

[Li] W.-C. W. Li, Newforms and functional equations, *Math. Ann.* **212** (1975), 285-315.

[LLY] W.-C. W. Li, L. Long, and Z. Yang, On Atkin-Swinnerton-Dyer congruence relations, *J. Number Theory*, to appear.

[Sc1] A. J. Scholl, Modular forms and de Rham cohomology; Atkin-Swinnerton-Dyer congruences, *Invent. Math.* **79** (1985), 49–77

[Sc2] A. J. Scholl, Fourier coefficients of Eisenstein series on noncongruence subgroups, *Math. Proc. Cambridge Philos. Soc.* **99** (1986), no. 1, 11–17.

[Sc3] A. J. Scholl, Modular forms on noncongruence subgroups. *Séminaire de Théorie des Nombres, Paris 1985-86*, 199-206, Progr. Math. 71, Birkhäuser Boston, boston, MA, 1987.

[Sc4] A. J. Scholl, The ℓ-adic representations attached to a certain non-congruence subgroup, *J. Reine Angew. Math.* **392** (1988), 1–15.

[Sc5] A. J. Scholl, The ℓ-adic representations attached to non-congruence subgroups II, preprint, 1993.

[Sc6] A. J. Scholl, On the Hecke algebra of a noncongruence subgroup,
Bull. London Math.Soc. **29** (1997), 395-399.

[Sc7] A. J. Scholl, On some ℓ-adic representations of $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ attached to noncongruence subgroups, preprint, 2003.

[Ser] J. P. Serre, Résumé de cours, 1984-5, Collège de France

[Sh1] G. Shimura, Sur les intégrales attachées aux formes automorphes, *J. Math. Soc. Japan* **11** (1959), 291–311.

[Sh2] G. Shimura, *Introduction to the arithmetic theory of automorphic functions*, Publications of the Mathematical Society of Japan, No. 11. Iwanami Shoten, Publishers, Tokyo, 1971, Kanô Memorial Lectures, No. 1.

[SW] C. M. Skinner and A. J. Wiles, Residually reducible representations and modular forms, *Inst. Haute. Etudes Sci. Publ. Math.* **89**, (1999), 5-126 (2000).

[TW] R. Taylor and A. Wiles, Ring-theoretic properties of certain Hecke algebras, *Ann. of Math.* **141** (1995), 553-572.

[Th] J. Thompson, Hecke operators and non congruence subgroups, in *Group Theory*, de Gruyter, New York, 1989, pp. 219-224.

[Wi] A. Wiles, Modular elliptic curves and Fermat's last theorem, *Ann. of Math.* **141** (1995), 443-551.

Wen-Ching Winnie Li
Department of Mathematics Pennsylvania State University, University Park, PA 16802 USA
E-mail:wli@math.psu.edu.

Ling Long
Department of Mathematics Iowa State University Ames, IA 50011 USA
E-mail: linglong@iastate.edu.

Zifeng Yang
Department of Mathematics Capital Normal University , Beijing 100037 P.R.China
E-mail:yangzf@mail.cnu.edu.cn.