11-30-2008

Do the Coefficients of a Modular Form Really "Encode Arithmetic Data"?

Ken McMurdy
Ramapo College of New Jersey, kmcmurdy@ramapo.edu

Hari Ravindran
Rose-Hulman Institute of Technology

Do the Coefficients of a Modular Form Really "Encode Arithmetic Data"?

Ken McMurdy and Hari Ravindran

Mathematical Sciences Technical Report Series
MSTR 08-01

November 30, 2008

Department of Mathematics
Rose-Hulman Institute of Technology
http://www.rose-hulman.edu/math

Fax (812)-877-8333 **Phone (812)-877-8193**

Do the Coefficients of a Modular Form Really "Encode Arithmetic Data"?

Ken McMurdy
Hari Ravindran

November 30, 2008

1 Introduction

Language and terminology are so critical to the understanding of modern mathematics that it is often difficult for even very good mathematicians from different fields to discuss their work in any detail. As a result, common phrases often evolve within each discipline which attempt to capture the flavor of some important idea while avoiding technicality and jargon. For example, when algebraic number theorists are asked why they are so interested in modular forms, it has become common to say with enthusiasm that the coefficients of a modular form "encode arithmetic data." If pressed further, one might go on to say that the modular form gives rise to a Galois representation (in some cases via an elliptic curve).

Unfortunately, experts may feel quite satisfied that the phrase does indeed convey its intended notion, while in reality very little of the meaning is conveyed to the broader mathematical community. Indeed, this was brought to our attention regarding the above phrase, "encode arithmetic data," when the results of this paper were originally presented. The objection was raised by one of the attendees that although he had heard the phrase before, it was still unclear to him that anything had been encoded (let alone arithmetic data). Thus, the phrase either conveyed no information or false information.

So this objection has provided the underlying philosophy of our paper. We aim to illustrate with a clear and concrete example how the coefficients of a modular form actually *do* encode arithmetic data, so that in the future this common phrase may carry greater meaning for a broader audience. In the process we also hope to concretely illustrate some of the various connections between modular forms, elliptic curves, and Galois representations, notions which have enjoyed considerable prominence in modern Number Theory, particularly due to their central role in the proof of Fermat's Last Theorem.

To be slightly more precise, we will verify a fundamental theorem of Shimura (given below as Theorem 4.1) for one particular modular form f by completely explicit means. In general, the theorem associates to a certain type of modular

form f a representation ρ_f of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ into GL_2 of some ring. Moreover, the theorem states that a great deal of information about ρ_f can be determined directly from the coefficients of f (indeed, enough to uniquely determine it), and this is precisely the basis on which we will claim that the coefficients of the modular form have "encoded" arithmetic data. Now, when the modular form has rational coefficients, as it will in our example, the representation can be seen as coming first from an associated elliptic curve. So after a brief review of some of the basic notions, an outline of our paper would be as follows. We begin with a specific modular form, associate to it an elliptic curve, compute the representation, and verify the theorem. We conclude with a few suggested problems for further exploration by the reader, and by elaborating briefly on some related results that provide a historical context for the paper.

2 Modular Curves and Modular Forms

Although modular forms can be defined and studied without any mention of modular curves, our work will center around the curves themselves, and will rely heavily on the "geometric" interpretation of (some) modular forms as functions or differentials on the curves. So we begin with a brief review of modular curves, closely following the development in [K, Ch. III, §1] (to which we refer the interested reader for more details).

Let $\mathbb{H}$ denote the complex upper half-plane, i.e. those complex numbers $a+bi$ for which $b > 0$, and let $\overline{\mathbb{H}} = \mathbb{H} \cup \mathbb{Q} \cup \{\infty\}$. Let $\Gamma = \mathrm{SL}_2(\mathbb{Z})$, the group of 2×2 matrices with integer coefficients and determinant 1, under multiplication. It is straightforward to show that Γ acts on $\overline{\mathbb{H}}$ by fractional linear transformations:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} (z) = \frac{az+b}{cz+d}.$$

Of course $\overline{\mathbb{H}}$ also then has an action by any subgroup of Γ. In particular, we will be most interested in the action of the subgroup, $\Gamma_0(N)$, consisting of those matrices in $\mathrm{SL}_2(\mathbb{Z})$ for which the lower left-hand entry is divisible by N. This action is sufficiently well-behaved that the quotient space, $\overline{\mathbb{H}}/\Gamma_0(N)$, inherits a complex analytic structure from $\mathbb{H}$, actually forming a compact Riemann surface (equivalently, a smooth projective curve over $\mathbb{C}$) which is called the "modular curve $X_0(N)$."

One way to visualize this classical construction of $X_0(N)$, for small N, is to first find an explicit fundamental domain for the group action. By this we mean a closed subset, $F \subseteq \mathbb{H}$, such that every equivalence class contains at least one point of F, and no two interior points of F are equivalent. Once a fundamental domain has been chosen, the modular curve is essentially obtained by gluing equivalent sections of the boundary of F together. The space is then compactified by including all the points which come from $\mathbb{Q} \cup \{\infty\}$ (smooth points which are called the "cusps" of the modular curve).

So how exactly does one find a fundamental domain? Typically one starts with the well-known (see [K, Ch. III, §1], for example) fundamental domain F

for all of Γ, given by

$$F = \{\ z \in \mathbb{H} \ \big|\ -\tfrac{1}{2} \leq \operatorname{Re} z \leq \tfrac{1}{2},\ |z| \geq 1\ \}.$$

Then a fundamental domain for $\Gamma_0(N)$ can be obtained by taking $\cup \gamma_i^{-1} F$, where $\{\gamma_i\}$ is any complete set of left coset representatives for $\Gamma/\Gamma_0(N)$. For example, $\Gamma_0(3)$ can be shown to have index 4 in $\mathrm{SL}_2(\mathbb{Z})$, with left coset representatives:

$$\alpha_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \alpha_1 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad \alpha_2 = \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}, \quad \alpha_3 = \begin{bmatrix} -1 & 1 \\ -1 & 0 \end{bmatrix}.$$

In Figure 1 below, we see the resulting fundamental domain for $\Gamma_0(3)$, letting F_i denote $\gamma_i^{-1} F$ for each i. In addition to the lines, $\operatorname{Re} z = \pm\frac{1}{2}$, and the circle, $|z| = 1$, the remaining curves pictured in the diagram are circles given by:

$$x^2 \pm 2x + y^2 = 0$$
$$x^2 \pm \tfrac{2}{3}x + y^2 = 0$$

for $z = x + yi$. When equivalent parts of the boundary are identified as indicated by the arrows, one arrives at a Riemann surface which is analytically isomorphic to $P^1(\mathbb{C})$. Note that the (compact) space, which we refer to as the modular curve $X_0(3)$, has precisely two cusps coming from ∞ and 0.

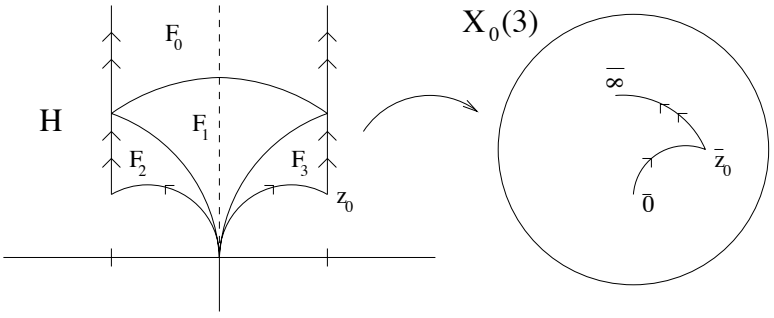

Figure 1: Classical Construction of $X_0(3)$

With a firm grasp of classical modular curves, consider now the following classical definition of a modular form.

Definition 2.1. A modular form of weight $k \in \mathbb{Z}$ for $\Gamma_0(N)$ is a meromorphic[1] function f on $\mathbb{H}$ such that:

$$f(\gamma z) = (cz + d)^k f(z), \quad \forall\ \gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \Gamma_0(N).$$

[1] This includes meromorphicity at the cusps, which we omit as a technical condition.

When $k = 0$, the condition that $f(\gamma z) = f(z)$ simply means that f passes to a well-defined function on the quotient space, i.e. f may be regarded as a function on the modular curve. Similarly, when $k = 2$, $f(z)\mathrm{d}z$ gives a well-defined differential on the modular curve which we denote by ω_f. For any other fixed k, the modular forms of weight k form an *invertible sheaf* on $X_0(N)$, but for this paper it will suffice to understand these two particular cases.

Now, regardless of N, it is clear that $\Gamma_0(N)$ will always contain the matrix γ for which $c = 0$ and $a = b = d = 1$. Thus, every modular form f satisfies $f(z + 1) = f(z)$. This implies that f is the pullback of some meromorphic function on a neighborhood of $q = 0$ by the map $q = e^{2\pi iz}$. The resulting Laurent series expansion of f in q (i.e. Fourier series expansion for f) is called the canonical q-expansion of f at infinity and usually denoted simply by $f(q)$. One should be a little careful here, however, as there is a subtle point about q-expansions. When $k = 0$, $f(q)$ really is the Fourier expansion of the function on $X_0(N)$ which corresponds to f. When $k = 2$, however, the differential on $X_0(N)$ which corresponds to f is actually $f(q)\frac{\mathrm{d}q}{q}$. So in particular, differentials which are holomorphic at the cusp ∞ appear to vanish in their canonical q-expansion. More generally, differentials which are holomorphic everywhere correspond to those weight 2 modular forms which appear to vanish at each cusp and hence are called *cusp forms* (see [K, Ch. III, §3] for details).

There are a variety of ways that one can obtain the q-expansions of modular forms. For this article we needed the q-expansions of all weight 2 cusp forms as well as enough weight 0 forms to generate the function field, for the specific modular curve, $X_0(26)$. For the weight 2 cusp forms (a two-dimensional vector space in this case), we found the following basis on William Stein's website.[2]

$$f_1 = q - q^2 + q^3 + q^4 - 3q^5 - q^6 + O(q^7)$$
$$f_2 = q + q^2 - 3q^3 + q^4 - q^5 - 3q^6 + O(q^7).$$

For an initial set of generating functions (i.e. weight 0 forms) on the modular curve, we chose the following two eta products:

$$t = \frac{\eta_2^2\eta_{13}^2}{\eta_{26}^2\eta_1^2} \qquad u = \frac{\eta_2^4\eta_{13}^2}{\eta_{26}^4\eta_1^2}.$$

Here we are using η_d as shorthand for the formal q-expansion given by

$$\eta_d = q^{\frac{d}{24}} \prod_{n=1}^{\infty}(1 - q^{nd}).$$

While η_d is not a modular form, the above eta products *are* legitimate functions, which can be verified by means of Ligozat's criterion (see [L, 3.2.1]).

[2]These are both Hecke newforms, which is a technical but crucial part of the hypothesis in Shimura's theorem. The website is http://modular.fas.harvard.edu/Tables.

3 Algebraic Curves and Elliptic Curves

Without going into great technical detail, it may be useful to include here a brief and intuitive review of the theory of algebraic curves over a field K. We recommend [Si, §II] as an excellent first introduction for newcomers to the field, and begin by paraphrasing Remark 2.5 from *Ibid.*

Remark 3.1. There is an equivalence of categories between smooth, projective curves defined over K (together with non-constant rational maps over K), and extensions F/K of transcendence degree 1 such that $F \cap \overline{K} = K$ (together with field injections fixing K).

This result provides a dictionary which translates between the geometric viewpoint (where most of the intuition comes from) and the algebraic viewpoint (where most of the calculations are done). It also explains the process by which we obtain our explicit equations for modular curves. In particular, we choose weight 0 modular forms which can be thought of as functions on the modular curve. Then, using the q-expansions, we find equations relating the functions. These functions and equations can then be thought of as generators for the field extension F described above. In other words, we find an "equation for the curve" by finding an equation which relates particular functions on the curve.

Now, there are also a few less philosophical results which bear mentioning. First of all, it may help to briefly review the theory of divisors of functions and differentials on a (smooth, projective) curve. For a point P on a curve C over $\overline{K}$, the functions which are holomorphic at P form a local ring which we denote by $\overline{K}[C]_P$. Any generator f of the maximal ideal, m_P, is called a "uniformizer at P". For all holomorphic functions f which are not identically zero, we then define $\mathrm{ord}_P(f)$ to be the maximal d for which $f \in m_P^d$. Intuitively, this is the order to which f vanishes at P. By extending in the obvious way to functions that have poles at P, we obtain a homomorphism

$$\mathrm{ord}_P : \overline{K}(C)^* \to \mathbb{Z}$$

where $\overline{K}(C)$ is the full function field of C. If one prefers to think analytically, functions at a particular point P have Laurent series expansions in any fixed uniformizer u. With this point of view, the ord of a function at P is simply the degree of the first term in its Laurent series expansion. For a differential ω on C, we may either define $\mathrm{ord}_P(\omega) = \mathrm{ord}_P(\omega/\mathrm{d}u)$, for any uniformizer u, or expand ω locally at P as $(u^d + \cdots)\mathrm{d}u$.

Let $\mathrm{Div}(C)$ denote the divisor group on C, i.e. the free abelian group generated by the set of points. Once ord at a point has been defined, it makes sense to define the divisor of a function f (or similarly a differential ω) by

$$\mathrm{Div}(f) = \sum_{P \in C} \mathrm{ord}_P(f) \cdot P \in \mathrm{Div}(C).$$

Thus, $f \mapsto \mathrm{Div}(f)$ defines a homomorphism from $\overline{K}(C)^*$ into $\mathrm{Div}(C)$. The kernel of this homomorphism consists precisely of the nonzero constant functions.

If we define the degree of a divisor to be the sum of its coefficients, the image of the homomorphism is contained in the subgroup $\mathrm{Div}^0(C)$ consisting of those divisors which have degree 0. The situation for differentials is quite different, however. The differentials which are holomorphic everywhere form a finite dimensional vector space whose dimension g is called the (geometric) genus of the curve. The divisor of any differential has degree $2g-2$.

Finally, in order to completely understand the theorem of Shimura, one should have some understanding of the Jacobian variety of a curve. As a group, the Jacobian is given by

$$J(C) = \mathrm{Div}^0(C)/\mathrm{Div}(\overline{K}(C)^*).$$

It is not immediate that this will always have the structure of an algebraic variety, but it does. Moreover, the dimension of $J(C)$ is always equal to the genus of C. The Jacobian of the modular curve $X_0(N)$ is denoted $J_0(N)$.

3.1 Elliptic Curves

Elliptic curves over K are smooth projective curves of genus 1 with a point defined over K. Assuming that the characteristic of K is not 2 or 3, such a curve always has an equation of the form

$$y^2 = x^3 + ax^2 + bx + c,$$

which is called a Weierstrass equation for the curve (the point at infinity is the K-rational point). In terms of these parameters, it is easy to check that the one dimensional space of holomorphic differentials is spanned by the differential, $\omega = dx/y$, which is both holomorphic and non-vanishing. As this differential is invariant under translation (by the group law, discussed below), it is often referred to as the "invariant differential."

The amazing fact about elliptic curves which makes them so useful for number theorists is that they have a natural group structure which can be given in terms of rational functions. Geometrically, we simply take the unique point at infinity to be the group identity, and then define $P + Q + R = \infty$ whenever P, Q, and R are collinear. There is one subtlety here, i.e. that a line which is tangent at P must be thought of as intersecting the curve twice at P (or 3 times at a point of inflection). Another way to think of the group structure on the elliptic curve is via its Jacobian. Since the genus of the elliptic curve E is 1, its Jacobian variety $J(E)$ again has dimension 1. The map from E to $J(E)$ which takes a point P to the class of the divisor $P - \infty$ can be shown to be an isomorphism between the two curves. Moreover, the group structure on $J(E)$ which is then passed along to E coincides with the one described above.

The connection between elliptic curves and 2-dimensional Galois representations comes from looking at torsion subgroups. For any elliptic curve E over an algebraically closed field, the torsion subgroup $E[N]$ is isomorphic to $\mathbb{Z}/N\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z}$, as long as N is relatively prime to the characteristic of K (see [Si, III, Cor 6.4]). Thus, if E is defined over K, we may let the Galois group of

$\overline{K}/K$ act on $E[N]$ to obtain a homomorphism from $\mathrm{Gal}(\overline{K}/K)$ into $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$. As a special case, when E is defined over $\mathbb{Q}$, one can take the inverse limit of the torsion subgroups $E[\ell^n]$, for ℓ some prime, to obtain a representation from $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ into $\mathrm{GL}_2(\mathbb{Z}_\ell)$. This construction of the so-called ℓ-adic Tate module of E is at the core of Shimura's theorem and will be discussed in far greater detail when it comes up in Section 7 below (or see [Si, III, §7]).

4 "Verifying" Shimura's Theorem

Now that we have reviewed some of the basics of modular forms and elliptic curves, we are able to state the theorem of Shimura whose verification in a special case is the central focus our paper. Many references for this theorem could be given (see [DI, §12.5], for example), but it is really a corollary of [Sh2, Theorem 1]. In order to most closely match our explicit example, we state the theorem first under the assumption that the modular form f has rational coefficients and hence corresponds to an elliptic curve over $\mathbb{Q}$. The more general case is addressed in Remark 4.2 below.

Theorem 4.1 (Shimura)**.** *Let f be a weight 2 newform[3] for $X_0(N)$ whose q-expansion at infinity is given by $f(q) = \sum_{n=1}^{\infty} a_n q^n$, with $a_n \in \mathbb{Q}$ for all n.*

(i) Then $X_0(N)$ surjects onto an elliptic curve E_f (defined over $\mathbb{Q}$), such that the invariant differential of E_f pulls back to a scalar multiple of ω_f.

(ii) For any prime ℓ, let $\rho_{f,\ell} : Gal(\overline{\mathbb{Q}}/\mathbb{Q}) \to GL_2(\mathbb{Z}_\ell)$ be the representation on the ℓ-adic Tate module of E_f. If p is any prime not dividing $N\ell$, then $\rho_{f,\ell}$ is unramified at p and for any Frobenius element σ_p at p we have

$$Tr\,\rho_{f,\ell}(\sigma_p) = a_p \quad \textit{and} \quad Det\,\rho_{f,\ell}(\sigma_p) = p.$$

We will partially verify this theorem in the case where $N = 26$, $\ell = 3$, and f is either of the two newforms listed above in Section 2 as f_1 and f_2. The first part of the theorem is completely verified when we explicitly demonstrate the two elliptic curve quotients of $X_0(26)$, E_{f_1} and E_{f_2}, in Section 6. For the second part, we are unable to list data for *all* primes p explicitly, and we can only approximate the 3-adic calculations. So we are content to handle a few carefully chosen primes, p, and to work up to (mod 9) precision only.

Remark 4.2. When f is defined over a more general number field K, the situation is only slightly more complicated. In this case, $J_0(N)$ has a direct factor A_f (over $\mathbb{Q}$) which is an abelian variety of dimension $[K : \mathbb{Q}]$. After extending scalars to $\mathbb{Q}_\ell$, the ℓ-adic Tate module of this A_f is a free module of rank 2 over $K \otimes \mathbb{Q}_\ell$. For any prime λ lying over ℓ in K, the injection of K_λ into $K \otimes \mathbb{Q}_\ell$ gives rise to a representation $\rho_{f,\lambda} : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2(K_\lambda)$ for which property (ii) from above still holds.

[3] Newforms are normalized cusp forms which are eigenvectors for the ring of Hecke operators, and which form a basis for the "new subspace." For full details see [DI, §6].

5 Explicit Model for $X_0(26)$

In order to verify Shimura's theorem, our first step is to find an explicit equation for the modular curve, $X_0(26)$. Recall that we have already specified two functions on the curve, given explicitly by the following eta products.

$$t = \frac{\eta_2^2\eta_{13}^2}{\eta_{26}^2\eta_1^2} \qquad u = \frac{\eta_2^4\eta_{13}^2}{\eta_{26}^4\eta_1^2}$$

To obtain an equation for the curve, we must find an algebraic equation relating these or some other parameters. By the q-expansion principal, this is equivalent to finding a relation between their q-expansions. For our final equation, we chose $x = t$ and $y = 2u - t^3 + 4t^2 + 4t - 1$. By comparing q-expansions, one sees that the following equation is satisfied.

$$y^2 = x^6 - 8x^5 + 8x^4 - 18x^3 + 8x^2 - 8x + 1 \tag{1}$$

Now, there are two important points to make about this equation. First of all, although we have an algebraic relation between two functions on $X_0(26)$, it is not immediate that the equation describes this curve. A priori, one only knows that the equation describes a quotient curve. Equivalently, the chosen parameters may only generate a subfield of the desired function field. One way to justify that we have the "whole thing," however, is to compare the genus of our quotient curve with the known genus of $X_0(26)$. Our equation describes a hyper-elliptic curve of genus 2, and using the well-known genus formula for $X_0(N)$ (see [Sh1, 1.40, 1.43], for example) we see that this is indeed the genus of the whole curve. Since any nontrivial quotient of a genus 2 curve would have strictly smaller genus by the Riemann-Hurwitz genus formula (see [Si, II, Thm 5.9]), the two curves must be the same.

The second point to make about our equation is that it is often not necessary to resort to brute force in finding relations between functions on a curve. The process is simplified, in the sense that the *form* of the equation can be determined in advance using the Riemann-Roch Theorem (see [Si, II, Thm 5.4]), if one knows where the functions under consideration have zeros and poles. Eta products are special in that they can only have zeros and/or poles at the cusps, and it is fairly straightforward to compute their orders by a variety of methods. We used families of Tate curves as in [M1], although classical methods can also be used. In this case we found that the divisors of our parameters were:

$$\begin{aligned} \mathrm{Div}(t) &= -(0) + (1/2) + (1/3) - (\infty) \\ \mathrm{Div}(u) &= 3(1/2) - 3(\infty). \end{aligned}$$

Moreover, the value of u at the cusp 0 is 13, so that both u and $t(u-13)$ are holomorphic outside of one point (the cusp ∞) where they had poles of order 3 and 4 respectively. For those familiar with Riemann-Roch Theorem, this suggested an equation relating functions inside $L(12\infty)$ which resulted in our initial model.

6 Elliptic Curve Quotients of $X_0(26)$

Now that we have an explicit equation for $X_0(26)$, we would like to find for each f the surjection ϕ_f onto an elliptic curve E_f which is guaranteed by the main theorem. If we were to follow Shimura's construction literally (as in the proof of [Sh2, Thm 1]), we would have to determine the appropriate ideals $\mathcal{U}$ of the Hecke algebra and take $E_f = J_0(26)/\mathcal{U}J_0(26)$. Subsequently we would compose the corresponding surjection with the canonical embedding of $X_0(26)$ into its Jacobian. There is a far more down-to-earth approach, however. Recall that f can be identified with a holomorphic differential ω_f on the modular curve. Shimura's construction is compatible with this identification in the sense that ω_f is the pullback by ϕ_f of the unique (up to scalar) holomorphic differential on E_f. Thus, using q-expansions to match f with an explicit holomorphic differential on our model for $X_0(26)$ reduces the problem of finding E_f to a straightforward algebraic curves exercise.

Our first step, then, is to find a basis for the holomorphic differentials on $X_0(26)$ in terms of the parameters from Equation (1). We leave it as an exercise to show that the vector space of holomorphic differentials is spanned by $\{\frac{dx}{y}, x\frac{dx}{y}\}$, and that the q-expansions of the corresponding weight 2 cusp forms are as follows.

$$\frac{dx}{y}(q) = -q^2 + 2q^3 - q^5 + q^6 - q^7 - q^8 + O(q^9)$$
$$x\frac{dx}{y}(q) = -q + q^3 - q^4 + 2q^5 + 2q^6 + O(q^9)$$

So just by comparing q-expansions we see that f_1 corresponds to the holomorphic differential $\omega_{f_1} = (1-x)\frac{dx}{y}$. Thus we know that $X_0(26)$ has an elliptic curve quotient $E_1 := E_{f_1}$ whose invariant differential pulls back to a scalar multiple of this differential. The only question is how to find it.

The key is to calculate the divisor of ω_{f_1}. While the divisor of the invariant differential on E_1 is of course 0, the divisor of ω_{f_1} is found to be $(1,4i)+(1,-4i)$. If one understands how divisors of functions and differentials behave with respect to maps between curves, it is possible to conclude quite a bit from this about the quotient map from $X_0(26)$ to E_1. In particular, it must be doubly ramified at these two points and unramified elsewhere (see [Si, pg. 28] and [Si, II, Prop. 4.3] for a precise explanation). Assuming the simplest explanation, i.e. that the quotient is of degree 2 and hence Galois, the problem thus reduces to finding an involution of $X_0(26)$ which fixes these two points. After a brief search, we found the involution:

$$\alpha_1(x,y) = (1/x, y/x^3).$$

In order to describe the quotient curve, we choose functions which are fixed by α, and then find a relation between them which follows from the equation for $X_0(26)$. For example, one could take $z_1 = x + 1/x$ and $w_1 = y/x + y/x^2$ and obtain the equation

$$w_1^2 = z_1^4 - 6z_1^3 - 11z_1^2 + 8z_1 - 4.$$

For the more familiar Weierstrass form, and our final equation for E_1, we make the further change of variables $x_1 = -2(w_1 - z_1^2 + 3z_1 + 10)$ and $y_1 = (4z_1 - 6)(w_1 - z_1^2 + 3z_1 + 10) + 52$ to obtain

$$E_1 : \quad y_1^2 = x_1^3 + 49x_1^2 + 728x_1 + 2704. \tag{2}$$

Analogously, by comparing q-expansions one finds that f_2 corresponds to the holomorphic differential $(-1-x)\frac{dx}{y}$. This has divisor $(-1, 2\sqrt{13})+(-1, -2\sqrt{13})$, which subsequently leads to the involution $\alpha_2(x, y) = (1/x, -y/x^3)$. Choosing parameters $z_2 = x + 1/x$ and $w_2 = y/x - y/x^2$ on $X_0(26)/\alpha_2$ we arrive at the equation,

$$w_2^2 = z_2^4 - 10z_2^3 + 21z_2^2 - 12z_2 + 4.$$

Again, we may make one final change of variables to obtain a Weierstrass equation, namely $x_2 = -2(w_2 - z_2^2 + 5z_2 + 2)$ and $y_2 = (4z_2 - 10)(w_2 - z_2^2 + 5z_2 + 2) + 32$. This results in the equation:

$$E_2 : \quad y_2^2 = x_2^3 + 33x_2^2 + 320x_2 + 1024 \tag{3}$$

Remark 6.1. Inside the function field of $X_0(26)$, the subfields corresponding to our two elliptic curve quotients intersect in $K(z)$ where $z = x+1/x$. Thinking of $K(z)$ as the function field of $\mathbf{P}^1$, we have decomposed $X_0(26)$ as a fiber product:

$$X_0(26) \cong E_1 \times_{\mathbf{P}^1} E_2$$

(where both projection maps have degree 2). As the Jacobian of $\mathbf{P}^1$ is trivial, this translates by functoriality of the Jacobian to an explicit splitting of $J_0(26)$ into the product of the two elliptic curves. Moreover, the splitting is essentially unique, since these two elliptic curves can be shown to be non-isogenous.

7 Associated ℓ-adic Representations

Now that we have good equations for the elliptic curves, E_1 and E_2, we are ready to compute the associated ℓ-adic representations by letting $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ act on the two ℓ-adic Tate modules. This sounds complicated, but it really isn't. Philosophically, the Tate module of an elliptic curve is constructed much like the ℓ-adic integers themselves. In particular, one can think of $\mathbb{Z}_\ell$ as the inverse limit of the following system of rings (with the usual reduction maps).

$$\cdots \to \mathbb{Z}/\ell^3\mathbb{Z} \to \mathbb{Z}/\ell^2\mathbb{Z} \to \mathbb{Z}/\ell\mathbb{Z} \to 0$$

So an element of $\mathbb{Z}_\ell$ is essentially nothing more than a consistent sequence of approximations in $\mathbb{Z}/\ell^n\mathbb{Z}$ for arbitrary n. Similarly, we think of the ℓ-adic Tate module of E as the inverse limit of the torsion subgroups $E[\ell^n]$, only with multiplication by ℓ providing the connecting homomorphisms. Each torsion subgroup, $E[\ell^n]$, is isomorphic to $\mathbb{Z}/\ell^n\mathbb{Z} \times \mathbb{Z}/\ell^n\mathbb{Z}$ and is acted upon *linearly* by

$\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ (basically because the addition law on E is given by polynomials). Thus we can construct a sequence of homomorphisms,

$$\rho_{E,n} : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{Aut}(E[\ell^n]) \cong \mathrm{GL}_2(\mathbb{Z}/\ell^n\mathbb{Z}).$$

The ℓ-adic representation $\rho_E : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2(\mathbb{Z}_\ell)$ is simply what we get if we paste together all of these approximations, $\rho_{E,n}$, in the usual ℓ-adic sense. The only real subtlety is in the "consistency" of the approximations. Since multiplication by ℓ is the homomorphism we consider between the torsion subgroups, the approximations will only (necessarily) satisfy

$$\rho_{E,n}(\sigma) \equiv \rho_{E,n-1}(\sigma) \pmod{\ell^{n-1}}$$

if we choose our basis for $E[\ell^{n-1}]$ to be ℓ times the basis for $E[\ell^n]$.

7.1 Representation Associated to $E[3]$

For brevity, we will treat only the 3-adic representation associated to $E := E_1$, explicitly computing the approximations coming from $E[3]$ and $E[9]$. So we begin by computing the 3-torsion subgroup of the elliptic curve E which is given by

$$y^2 = x^3 + 49x^2 + 728x + 2704.$$

The easiest way to do this is probably to use the well-known duplication formula (see [Si, III, 2.3]), along with the facts that $2P = -P$ whenever $3P = 0$ and $x(-P) = x(P)$. This yields the following.

$$\begin{aligned} \frac{x^4 - 1456x^2 - 21632x}{4x^3 + 196x^2 + 2912x + 10816} &= x \\ x(3x + 52)(x^2 + 48x + 624) &= 0 \end{aligned}$$

Note that the four x values result in eight points on the curve, which makes sense as they should form a group isomorphic to $\mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$ (along with the infinite point as the identity). Note also that we must specify a basis for $E[3]$ before the representation can literally take values in $\mathrm{GL}_2(\mathbb{Z}/3\mathbb{Z})$. For our basis, we choose the ordered points, $P = (0, 52)$ and $Q = (-52/3, 104\sqrt{-3}/9)$, which results in the $\mathbb{Z}/3\mathbb{Z}$-module structure explicitly shown below in Table 1.

Now it's pretty easy to see from the group table how Galois acts on $E[3]$. If $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ happens to fix the square root of -3, well then it acts trivially on all of $E[3]$. On the other hand, if $\sigma(\sqrt{-3}) = -\sqrt{-3}$ (the only other option), then σ fixes P but takes Q to $-Q$. Remember that the action is linear in any case. So we can describe $\rho_{E,1}$ explicitly by

$$\rho_{E,1}(\sigma) = \begin{cases} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, & \text{if } \sigma(\sqrt{-3}) = \sqrt{-3} \\ \\ \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, & \text{if } \sigma(\sqrt{-3}) = -\sqrt{-3}. \end{cases}$$

$+$	∞	$(-\frac{52}{3}, \frac{104\sqrt{-3}}{9})$	$(-\frac{52}{3}, -\frac{104\sqrt{-3}}{9})$
∞	∞	$(-\frac{52}{3}, \frac{104\sqrt{-3}}{9})$	$(-\frac{52}{3}, -\frac{104\sqrt{-3}}{9})$
$(0, 52)$	$(0, 52)$	$(-24 - 4\sqrt{-3}, 28 - 4\sqrt{-3})$	$(-24 + 4\sqrt{-3}, 28 + 4\sqrt{-3})$
$(0, -52)$	$(0, -52)$	$(-24 + 4\sqrt{-3}, -28 - 4\sqrt{-3})$	$(-24 - 4\sqrt{-3}, -28 + 4\sqrt{-3})$

Table 1: Isomorphism of $E[3]$ with $\mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$

The key question remains, however, "Does this agree with what Shimura's theorem predicted?" Well the answer is a very satisfying, "Yes!"

To verify the theorem, we consider only those primes not equal to 2, 3, or 13. For any such p, we basically reduce the whole group table mod p and then ask how the Frobenius automorphism, defined by $\sigma_p(a) = a^p$ for $a \in \overline{\mathbb{F}}_p$, acts. Comparing with the coefficients of our original modular form f_1, we should see in each case that

$$\operatorname{Tr} \rho_{E,1}(\sigma_p) \equiv a_p, \quad \operatorname{Det} \rho_{E,1}(\sigma_p) \equiv p \pmod 3.$$

For example, there is no $\sqrt{-3}$ in $\mathbb{F}_{11}$, and hence σ_{11} acts in the only allowable nontrivial way with determinant -1 and trace 0. So this checks, as $11 \equiv -1$ and $a_{11} = 6 \equiv 0 \bmod 3$. For the other type of example, consider $p = 19$. Since $\sqrt{-3}$ reduces to either ± 4 in $\mathbb{F}_{19}$, the whole group table is defined over $\mathbb{F}_{19}$. Therefore Frobenius acts trivially, with determinant 1 and trace -1. Well this also agrees with Shimura, since $19 \equiv 1$ and $a_{19} = 2 \equiv -1 \bmod 3$. Every other prime will of course look like one of these two. In fact, quadratic reciprocity makes it possible to sum up what happens for all p. There is a square root of -3 in $\mathbb{F}_p$ (for $p \geq 5$) precisely when $p \equiv 1 \bmod 3$. Hence, these are the cases for which $\rho_{E,1}(\sigma_p)$ has determinant 1 and trace -1. In such cases, the theorem then tells us that $p \equiv 1$ (immediate) and $a_p \equiv -1 \bmod 3$. Likewise, when $p \equiv -1 \bmod 3$ (and $\sqrt{-3}$ generates $\mathbb{F}_{p^2}$), the theorem implies that $p \equiv -1$ (immediate) and $a_p \equiv 0 \bmod 3$. So verification of the theorem comes down to checking that $a_p \equiv -1$ when $p \equiv 1 \bmod 3$, and $a_p \equiv 0$ whenever $p \equiv -1 \bmod 3$. For convenient checking of the first several a_p values, we collect them below in Table 2.

p	5	7	11	13	17	19	23	29	31	37	41	43	47
a_p	-3	-1	6	1	-3	2	0	6	-4	-7	0	-1	3

Table 2: Values of a_p for the modular form f_1

7.2 Representation Associated to $E[9]$

Having completed the previous calculation as a sort of "warm-up," we would like to now compute the representation $\rho_{E,2}$ and verify that Shimura's theorem still predicts the correct traces and determinants for Frobenius elements (now working mod 9). Of course, it is not practical to show the entire group table for $E[9]$, whose 81 elements lie in an algebraic extension of $\mathbb{Q}$ of degree 162. We don't really *need* to do this, though, to verify the theorem. By linearity, all we need to do is choose a specific basis for $E[9]$, and then see what various Frobenius automorphisms do to these basis elements.

Remember that we want to see consistency between $\rho_{E,1}$ and $\rho_{E,2}$, and therefore we must in addition choose the basis so that $3P_1 = P$ and $3Q_1 = Q$. In other words, we need to "divide" the points P and Q by 3. The easiest way to do this is probably just to derive a "triplication formula" on the curve, analogous to the duplication formula we used earlier. An inductive process for generating all such formulas is outlined quite explicitly in [Si, Ex. 3.7]. Alternatively, one can simply compute $3R = 2R + R$ for $R = (x, y)$ directly, using the addition and duplication formulas in [Si]. Either way, one quickly arrives at $x(3R) = \Phi_3(x(R))/(\Psi_3(x(R)))^2$, where

$$\begin{aligned}\Phi_3(x) = x^9 &- 8736x^7 - 544960x^6 - 12719616x^5 \\ &- 47244288x^4 + 3437584384x^3 + 68787683328x^2 \\ &+ 510994219008x + 1265319018496\end{aligned}$$

$$\Psi_3(x) = x(3x + 52)(x^2 + 48x + 624).$$

Using the formula, we now choose basis vectors for $E[9]$ in the following way. First we choose roots of $\Phi_3(x)/(\Psi_3(x))^2 = 0$ and $\Phi_3(x)/(\Psi_3(x))^2 = -52/3$, which are (choices for) the x coordinates of P_1 and Q_1. Then trial and error with $3P_1 = P$ and $3Q_1 = Q$ gives us the correct choice (2 options) for the y coordinates. Now, this could be done completely globally, i.e. over a finite extension of $\mathbb{Q}$. As we are only looking to verify the action of Frobenius automorphisms on $E[9]$, however, there is no reason not to check prime-by-prime and do all of our arithmetic over finite fields. So for each prime p, we begin by specifying (via minimal polynomial) the finite field $\mathbb{F}_p[\gamma]$ over which the coordinates of P_1 and Q_1 are defined. Then we apply σ_p to the coordinates of P_1 and Q_1, and determine the unique coefficients $a, b, c, d \in \mathbb{Z}/9\mathbb{Z}$ such that $\sigma_p(P_1) = aP_1 + bQ_1$ and $\sigma_p(Q_1) = cP_1 + dQ_1$. This is done by computing all linear combinations of P_1 and Q_1 over $\mathbb{Z}/9\mathbb{Z}$ and comparing, but we note that one already knows what these coefficients are mod 3 by compatibility with $\rho_{E,1}$ and our previously computed $\sigma_p(P)$ and $\sigma_p(Q)$. So this narrows down the search considerably. Thus, with respect to the basis $\{P_1, Q_1\}$, we will have found

$$\rho_{E,2}(\sigma_p) = \begin{bmatrix} a & c \\ b & d \end{bmatrix}.$$

For example, when $p = 17$ we may define P_1 and Q_1 over the field $\mathbb{F}_{17}[\gamma]$

where $f_{min}(\gamma) = \gamma^6 + 2\gamma^4 + 10\gamma^2 + 3\gamma + 3$. In particular, we may take

$$P_1 = (4\gamma^4 + 16\gamma^3 + 8\gamma^2 + \gamma + 10, 13\gamma^5 + 12\gamma^4 + 12\gamma^3 + 14\gamma^2 + 5\gamma + 12)$$
$$Q_1 = (\gamma^5 + 7\gamma^4 + 13\gamma^3 + 10\gamma^2 + 9\gamma + 7, 5\gamma^5 + 9\gamma^4 + \gamma^3 + 2\gamma^2 + 10\gamma + 8).$$

Note that it is straightforward to check that $3P_1$ and $3Q_1$ do indeed equal P and Q respectively[4]. Now to understand how σ_{17} acts on $E[9]$, we raise the coefficients of P_1 and Q_1 to the 17th power, and compare with linear combinations of P_1 and Q_1 over $\mathbb{Z}/9\mathbb{Z}$. Doing this, we find that $\sigma_{17}(P_1) = 4P_1 + 3Q_1$ and $\sigma_{17}(Q_1) = 0P_1 + 2Q_1$. In other words, we find that

$$\rho_{E,2}(\sigma_{17}) = \begin{bmatrix} 4 & 0 \\ 3 & 2 \end{bmatrix}.$$

Does this agree with Shimura? According to the main theorem, we should have $\mathrm{Det}\rho_{E,2}(\sigma_{17}) \equiv 17 \pmod 9$ and $\mathrm{Tr}\rho_{E,2}(\sigma_{17}) \equiv a_{17} \pmod 9$. These clearly both check, as $a_{17} = -3$ from Table 2. Following this same algorithm, we have also computed $\rho_{E,2}(\sigma_p)$ for the primes $p = 5$, 7, 11, and 19, and this data is summarized in Table 3 below. We invite the reader to verify that the trace and determinant agree with the values predicted by the main theorem in each case!

p	$f_{min}(\gamma)$	P_1	$aP_1 + bQ_1$
		Q_1	$cP_1 + dQ_1$
5	$\gamma^6 + \gamma^4 + 4\gamma^3 + \gamma^2 + 2$	$(4, 3)$	$1P_1 + 0Q_1$
		$(4\gamma^4 + \gamma^3 + 2\gamma^2 + 4\gamma, \gamma^5 + 3\gamma^4 + 4\gamma^3 + 2\gamma^2 + 3\gamma + 4)$	$3P_1 + 5Q_1$
7	$\gamma^3 + 6\gamma^2 + 4$	$(4\gamma + 2, 5\gamma^2 + 2\gamma + 2)$	$4P_1 + 0Q_1$
		$(3\gamma, \gamma^2 + 5\gamma + 6)$	$3P_1 + 4Q_1$
11	$\gamma^6 + 3\gamma^4 + 4\gamma^3 + 6\gamma^2 + 7\gamma + 2$	$(7\gamma^4 + 7\gamma^3 + 10\gamma^2 + 4\gamma + 10, \gamma^5 + 8\gamma^2 + 3\gamma)$	$7P_1 + 0Q_1$
		$(9, 3\gamma^5 + 7\gamma^4 + 3\gamma^3 + 6\gamma^2 + 9\gamma + 1)$	$0P_1 + 8Q_1$
19	$\gamma^3 + 4\gamma + 17$	$(15\gamma^2 + 16, 8\gamma^2 + 18\gamma + 8)$	$4P_1 + 6Q_1$
		$(18\gamma + 7, 4\gamma^2 + 9\gamma + 9)$	$0P_1 + 7Q_1$

Table 3: Action of σ_p on $E[9]$ for various p

[4]The explicit calculations on elliptic curves for this section will be omitted for brevity, but were performed using SAGE. The complete session is available for download on the first author's website.

8 Conclusion

For our main conclusion, we would like to argue on the basis of the preceding example that the coefficients of a modular form really do "encode arithmetic data." In very loose terms, a code is some sort of transformation which can be applied to data. If it's a good one, it should be fairly easy to apply in the forward direction, while attempting to go in the reverse direction should be very difficult without some sort of key. With the right point of view, this is precisely what is happening in Shimura's theorem.

In particular, the analogy works best if we think of the Galois representation as the original data, the traces and determinants of all Frobenius elements as the encoded data, and the actual modular form as the key. Remember that in practice a Galois representation looks like an underlying vector space (or module), with a basis that is defined over some extension of $\mathbb{Q}$, so that Galois acts linearly. For our first approximation, this vector space was two dimensional over $\mathbb{F}_3$, and the basis $\langle P, Q\rangle$ was defined over $\mathbb{Q}(\sqrt{-3})$. The second approximation was two dimensional over $\mathbb{Z}/9\mathbb{Z}$, and the basis $\langle P_1, Q_1\rangle$ was defined over a quadratic extension of the splitting field of $\phi_3(x) + \frac{53}{3}(\psi_3(x))^2$. Given one of these gadgets, it is of course very straightforward to generate the traces and determinants of Frobenius elements (precisely as we did in generating Table 3). Moreover, it follows from the Cebotarev Density Theorem that the representation is completely determined by these traces and determinants, although determining the representation from this information alone is very difficult. Thus it makes perfect sense to think of the traces and determinants as the encoded data, and by Shimura's theorem this is precisely what one sees in the *coefficients* of a modular form.

Interestingly, though, the modular form itself is also in some sense the "key" to unlocking the code. With only the traces and determinants of a two dimensional Galois representation, it would indeed be difficult to work backwards and come up with a matching module with explicit Galois action. On the other hand, if we know that the representation is *modular*, and we know the level, the problem boils down to the geometric problem of computing torsion points in the correct modular abelian variety, or in our case torsion points in an elliptic curve quotient of $X_0(N)$. Thus, in one fell swoop, Shimura's theorem has not only provided us with an extremely rich family of encoded Galois representations, but also the keys to unraveling the codes. Philosophically, this begins to explain why modular forms have been such exciting objects of study for so many number theorists!

8.1 Suggested Further Exploration

For the curious reader, there are a number of projects which would undoubtedly strengthen the understanding gained from this paper. For example, the easiest thing to do would be to simply stay with $X_0(26)$ and approximate the ℓ-adic representations associated to f_1 or f_2 for various other primes ℓ (we only did $\ell = 3$ and f_1).

For a similar example taken from a different modular curve, we suggest looking at $N = 49$ and then $N = 50$. The first curve already has genus 1 (so no quotient is necessary), and an explicit equation is given in [M2, §2]. The second has genus 2, and the weight 2 cusp forms are spanned by two newforms that are defined over $\mathbb{Q}$. So this example is extremely similar to our $X_0(26)$ example, with the Jacobian of $X_0(50)$ splitting into the product of two elliptic curves. Also, an explicit model for $X_0(50)$ can be easily derived using the equations for $X_0(25)$ given in [M2, §4] and the additional function $t_2 = (\eta_1/\eta_2)^{24}$ which satisfies $j = (t_2 + 256)^3/t_2^2$.

For a slightly more complicated example, we suggest looking at $X_0(35)$. An explicit model can be obtained by crossing $X_0(7)$ and $X_0(5)$ over the j-line using equations from [M2, §2, 4]. The curve has genus 3, and the weight 2 cusp forms are spanned by one newform defined over $\mathbb{Q}$ and two Galois conjugate newforms defined over a quadratic extension K of $\mathbb{Q}$. Thus, the Jacobian of $X_0(35)$ splits into an elliptic curve and an abelian surface. Computing the representation associated to the unique newform over $\mathbb{Q}$ would again be similar to the previous examples. For both conjugates over K, however, we would obtain just one representation by computing the ℓ-adic Tate module of the abelian surface and viewing it as a two dimensional $\mathcal{O}_K \otimes \mathbb{Z}_\ell$-module via the action of the Hecke algebra (as described in Remark 4.2).

Finally, in order to appreciate the encoding principle, it might be interesting to choose a modular form f and then attempt to construct the associated Galois representation using **only** the traces and determinants of Frobenius elements as determined by the a_p values. For example, using the q-expansion for our f_1, one should be able to reduce all of the a_p values mod 3 and then come up with something like Table 1 through pure algebraic number theory means (and no elliptic curves). This is essentially an explicit Cebotarev density problem for which various algorithms exist.

8.2 Connections

Another goal in writing this paper has been to assist in making a sizable chunk of modern number theory, not just a few specific results, more accessible to a wider range of mathematicians. To this end, we would now like to mention a few of the ways in which the results of this paper can serve as a bridge to some of the other results and areas of research in modern number theory.

A natural first connection to make is to the proof of Fermat's Last Theorem. We briefly outline that connection here, and recommend the reader to [CSS] as an excellent reference for more details. Recall that in Section 6, we showed explicitly that the elliptic curves E_1 and E_2 were "modular," in the sense that they were quotients of the modular curve $X_0(26)$. One of the key ingredients in the proof of Fermat's Last Theorem, the Shimura-Taniyama-Weil conjecture, states that in fact *all* elliptic curves defined over $\mathbb{Q}$ are modular in this same sense (quotients of $X_0(N)$ for varying N). Fermat's Last Theorem was proven when Andrew Wiles was able to establish the conjecture for all semi-stable elliptic

curves over $\mathbb{Q}$,[5] essentially completing a proof by contradiction. In particular, any counterexample to Fermat was known to imply the existence of a semi-stable elliptic curve whose associated ℓ-adic Galois representation would have certain very special properties. By a theorem of Ken Ribet, any such elliptic curve could not possibly be modular. Hence, after Wiles, no such counterexample to Fermat could exist.

Another natural and understandable connection to make is to the recent proof of Serre's Conjecture by Khare and Wintenberger. In Theorem 4.1, we saw a way to attach to a modular form f a representation of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ into GL_2 of some ℓ-adic ring. Such a representation can be reduced to obtain a representation into $\mathrm{GL}_2(\overline{\mathbb{F}}_\ell)$, and quite a bit of work has been done to understand how modular forms of different weights and levels can give rise to the same mod ℓ representation. Serre's Conjecture (now a theorem of Khare and Wintenberger) could be seen as a converse of sorts. It states that in fact all sufficiently nice two-dimensional mod ℓ Galois representations arise in this manner.

Finally, the concept of a modular form and associated Galois representation has been generalized substantially in a variety of ways. Overconvergent modular forms, for example, can be seen as sections of an invertible sheaf over a rigid-analytic subspace of a modular curve over $\mathbb{Q}_p$. Hilbert modular forms are functions on the m-fold product of complex upper half-planes which transform nicely under the action of a particular discrete subgroup of $\mathrm{GL}_2(F)$, where F is a totally real number field of degree m. Similarly, Siegel modular forms are functions on the "Siegel upper half-space" which (almost) respect the action of a certain subgroup of $\mathrm{GL}_{2g}(\mathbb{Z})$. Hilbert and Siegel modular forms (along with the classical ones) are both examples of Automorphic Forms, which give rise to what are called Automorphic Representations. The Langlands Program is an attempt to understand and classify a broad class of representations which in some sense include these, but with an emphasis on the algebraic groups rather than concrete individual constructions. For an introduction to the Langlands Program which shows explicitly how classical and Siegel modular forms fit into Langlands' theory, we recommend [G].

References

[CSS] G. Cornell, J. Silverman, G. Stevens, (editors) *Modular Forms and Fermat's Last Theorem*, Papers from the Instructional Conference on Number Theory and Arithmetic Geometry held at Boston University, Aug. 9–18, 1995, Springer-Verlag (1997).

[DI] F. Diamond, J. Im, Modular Forms and Modular Curves, Canadian Math. Society Conference Proceedings **17** (1995), 39–133.

[DS] F. Diamond, J. Shurman, *A First Course in Modular Forms*, Graduate Texts in Mathematics **228**, Springer-Verlag (2005).

[5] The conjecture has since been fully proven by Breuil, Conrad, Diamond, and Taylor.

[G] S. Gelbart, An Elementary Introduction to the Langlands program, Bull. Amer. Math. Soc. (N. S.) **10** (1984), no. 2, 177–219.

[K] N. Koblitz, *Introduction to Elliptic Curves and Modular Forms*, Graduate Texts in Mathematics **97**, Springer-Verlag (1984).

[L] G. Ligozat, *Courbes modulaires de genre* 1, Bull. Soc. Math. France, Mém. **43** (1975).

[M1] K. McMurdy, Eta Products for Modular Curves, Online Proceedings from the Workshop on Computations with Modular Forms, Heilbronn Institute, Bristol (2008). http://www.uni-due.de/ hx0037/CMF/

[M2] K. McMurdy, Explicit parameterizations of ordinary and supersingular regions of $X_0(p^n)$, Modular Curves and Abelian Varieties (Barcelona, 2002), 165–179, Prog. Math. **224** (2004).

[Sh1] G. Shimura, *Introduction to the Arithmetic Theory of Automorphic Functions*, Iwanami Shoten and Princeton Univ. Press, Princeton, 1971.

[Sh2] G. Shimura, *On the factors of the jacobian variety of a modular function field*, J. Math. Soc. Japan **25** (1973), no. 3, 523–544.

[Si] J. Silverman, *The Arithmetic of Elliptic Curves*, Graduate Texts in Mathematics **106**, Springer-Verlag (1986).