ELLIPTIC CURVES, SERRE'S CONJECTURE AND FERMAT'S LAST THEOREM

KIRTI JOSHI

Contents

Date: Version of Joshi 5/13/2000 9:35:17 AM.

1991 *Mathematics Subject Classification.* 14J20, 14C25.

Key words and phrases. elliptic curves, serre's conjecture, galois representations, fermat's last theorem, modular forms, modular representations, frey curves.

1. Introduction

In these lectures I want to explain a circle of ideas introduced about 15 years ago which led to the proof of Fermat's Last Theorem at the hands of Taylor and Wiles (see [49], [46]) and the subsequent refinement of these ideas at the hands of Breuil, Conrad, Diamond and Taylor led to a proof of the Shimura-Taniyama-Weil conjecture (see [3]). Needless to say that these two results rest on the work of a number of other mathematicians as well: H. Hida, B. Mazur, K. Ribet, F. Diamond, B. Edixhoven, P. Deligne, J.-P. Serre and many others (see [4], [5], [7], [8], [11], [17], [14], [15], [19], [24], [32], [33], [34], [37], [42], [46], [49], and references at the end of these articles)

The main conjecture which formed the backdrop of these developments is a conjecture of Serre (see [41]). Despite the developments which have taken place, this conjecture of Serre still remains intractable at the moment. In [41] Serre showed that his conjecture together with the observation of G. Frey (see [20], [21]) led to a proof of Fermat's Last Theorem.

In these lectures I will essentially outline a proof of this assertion that Serre's conjecture implies Fermat's last Theorem. This article, needless to say, is completely based on Serre's paper [41] and is meant to serve as an introduction to the circle of ideas introduced in Serre's paper and is, by no means, a substitute for it. I have attempted to keep this article as self contained as possible. However it is impossible to prove all the results or to develop the theory of Galois representations in any reasonable depth or detail in article of this length. A more comprehensive account of the subject can be found in [9] or [16] and Serre's book [38] is a classic introduction to the subject of Galois representations.

2. Some History

The term elliptic curves is of relatively recent vintage, but the fundamental objects which lead to these curves have been around for a long time. Euler and Legendre studied the following kind of complex integrals:

$$\int_z^\infty \frac{dx}{((1-x^2)(1-k^2x^2))^{1/2}} \text{ and } \int_z^\infty \frac{dx}{(4x^3-g_2x-g_3)^{1/2}} \tag{2.1}$$

where k, g_2, g_3 are complex numbers. Integrals like these arise naturally while calculating the length of the arc of an ellipse and hence these integrals were

called *elliptic integrals*. Functions which arise in inverting these integrals were called elliptic functions. These integrals naturally live on curves whose equations look like

$$y^2 = f(x) \tag{2.2}$$

where $f(x)$ is a polynomial in x of degree three or four with distinct roots and complex coefficients. Such curves are called *elliptic curves*.

3. Cubic Curves

Let K be a field. Consider curves defined by homogeneous polynomials of the form

$$Y^2Z + a_1XYZ + a_3YZ^2 = X^3 + a_2X^2Z + a_4XZ^2 + a_6Z^3 \tag{3.1}$$

where the coefficients $a_i \in K$. Observe that as the polynomial is homogeneous if (x_0, y_0, z_0) is a solution to (3.1) then so is $(\lambda x_0, \lambda y_0, \lambda z_0)$ for any constant $\lambda \neq 0$. So we discard the *trivial solution* $(0, 0, 0)$, and identify solutions which are non-zero multiples of one another. In other words we will work with *homogeneous* or *projective* coordinates X, Y, Z and hence forth identify the solution (x_0, y_0, z_0) with $(\lambda x_0, \lambda y_0, \lambda z_0)$ for any $\lambda \neq 0$.

Also note that (3.1) has another obvious solution $O = (0, 1, 0)$ and it is easy to check that this the only non-trivial solution with $z = 0$. We will call O the *point at infinity* on the curve.

If (x_0, y_0, z_0) is a solution with $z_0 \neq 0$ then we can scale the solution by $1/z_0$ to get another solution $(x_0/z_0, y_0/z_0, 1)$. Thus we see that, except O, any other solution of (3.1) can be taken to be of the form $(x_1, y_1, 1)$ for some (x_1, y_1) and that such solutions are points on the dehomogenised form of the equation

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6 \tag{3.2}$$

where we have written $y = Y/Z, x = X/Z$. For simplicity we will always write the equation (3.1) in its dehomogenous form (3.2) remembering the extra point O. We define several important quantities associated to the curve:

$$b_2 = a_1^2 + 4a_2 \tag{3.3}$$
$$b_4 = 2a_4 + a_1a_3 \tag{3.4}$$
$$b_6 = a_3^2 + 4a_6 \tag{3.5}$$
$$b_8 = a_1^2a_6 + 4a_2a_6 - a_1a_3a_4 + a_2a_3^3 - a_4^2 \tag{3.6}$$
$$c_4 = b_2^2 - 24b_4 \tag{3.7}$$
$$c_6 = b_2^2 + 36b_2b_4 - 216b_6 \tag{3.8}$$
$$\Delta = -b_2^2b_8 - 8b_4^3 - 27b_8^2 + 9b_2b_4b_6 \tag{3.9}$$
$$j = c_4^3/\Delta \tag{3.10}$$

We call Δ the *discriminant* of the curve.

Definition 3.11. An elliptic curve E/K is a curve given by

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6 \tag{3.12}$$

where $a_i \in K$ and provided that the discriminant $\Delta = \Delta_E \neq 0$.

Definition 3.13. A solution (x_0, y_0, z_0) of (3.11) will be called a point on the curve defined by the equation and if $x_0, y_0, z_0 \in K$ we will say that (x_0, y_0, z_0) is a K-rational point, or when there is no cause for confusion, simply a rational point of E.

Example 3.14. For example $y^2 = x^3 - x$ is an elliptic curve over any field K in which $\Delta_E = -64 \neq 0$.

Example 3.15. Let E be defined by $y^2 = x^3 + x + t$. Then we can easily calculate the quantities defined above: $b_2 = 0$, $b_4 = 2$, $b_6 = 4t$, $b_8 = -1$, $c_4 = -48$, $c_6 = -864t$, $\Delta = -432t^2 - 64$, $j = -6192/(-27t^2 - 4)$. Hence this equation defines an elliptic curve if and only if $-432t^2 - 64 \neq 0$.

So far we have not placed any restrictions over K, but under additional assumptions on K we can simplify the equation of any elliptic curve considerably. For instance if $char(K) \neq 2$ we can replace y by $\frac{1}{2}(y - a_1x - a_3)$ to get

$$y^2 = 4x^3 + b_2x^2 + 2b_4x + b_6 \tag{3.16}$$

And if $char(K) \neq 2, 3$ then we can replace (x, y) by $(\frac{x-3b_2}{36}, \frac{y}{216})$ to get

$$y^2 = x^3 - 27c_4x - 54c_6 \tag{3.17}$$

One gets the additional relations $4b_8 = b_2b_6 - b_4^2$ and $1728\Delta = c_4^3 - c_6^2$.

Remark 3.18. The only transformations of (3.11) which preserve the form of the equation are of the following kind:

$$x = u^2x' + r \tag{3.19}$$

$$y = u^3y' + u^2sx' + t \tag{3.20}$$

with $u, r, s, t \in K$ and $u \neq 0$. I leave it to you to check that under these substitutions the new value of the new discriminant is $\Delta' = u^{-12}\Delta$ and $j' = j$. Thus the function j of the coefficients a_1, a_2, a_3, a_4, a_6 is an invariant of the curve and is called the *j-invariant of the elliptic curve.*

4. Singularities

Suppose E is given by

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6 \tag{4.1}$$

but is not an elliptic curve, i.e., $\Delta_E = 0$. In this situation we will say that the curve is *singular* and we want to describe the geometric possibilities for E in this situation. Let

$$f(x, y) = y^2 + a_1xy + a_3y - (x^3 + a_2x^2 + a_4x + a_6) \tag{4.2}$$

Then it is not very difficult to show that $\Delta_E = 0$ if and only of

$$\frac{\partial}{\partial x}f(x, y) = \frac{\partial}{\partial y}f(x, y) = 0 \tag{4.3}$$

has a simultaneous solution, say (x_0, y_0). Then we can use Taylor series to write $f(x, y)$ as

$$f(x,y)-f(x_0,y_0) = [(y-y_0)-\alpha(x-x_0)][(y-y_0)-\beta(x-x_0)]-(x-x_0)^3 \tag{4.4}$$

where, $\alpha, \beta \in \bar{K}$, where $\bar{K}$ is an algebraic closure of K.

Remark 4.5. In general α, β may not be in K.

Definition 4.6. Let the notation be as above. If $\alpha \neq \beta$ we say that $P = (x_0, y_0)$ is a node and in this case the lines

$$y - y_0 = \alpha(x - x_0) \tag{4.7}$$

$$y - y_0 = \beta(x - x_0) \tag{4.8}$$

are two tangents to the curve E at (x_0, y_0).

If $\alpha = \beta$ we say that E has a cusp at $P = (x_0, y_0)$ and the line $y - y_0 = \alpha(x - x_0)$ is a multiple tangent to E at $P = (x_0, y_0)$.

Example 4.9. The curve $y^2 = x^3 + x^2$ is singular with a node at the point $(0, 0,)$.

Example 4.10. The curve $y^2 = x^3$ has a cusp at $(0, 0)$.

5. Group Law

Elliptic curves have many remarkable properties, and one of the important properties is that the set of points on an elliptic curve forms a group. I will briefly describe the group law. The group law is best described geometrically but it also admits an algebraic description.

Suppose E/K is an elliptic curve. Let L/K be any field extension. We will write $E(L)$ for the set of L-rational points of E. Note that $O \in E(L)$ for any L/K and so $E(L)$ is a non-empty set. If $P, Q \in E(L)$ are two points then the line joining P, Q intersects the curve in a third point. A little bit of algebra shows that this new point also has coordinates in L, i.e., it is also an L-rational point. This basic geometric fact underlies the group law on the elliptic curve. Let us denote this third point by R. Then we join O and R by a line. This line also intersects the curve again in a point $R' \in E(L)$. We declare that $P + Q = R'$. If $P = Q$ we take the line to be the tangent line to the curve at P. It is not hard to check that this makes the set $E(L)$, for any extension L/K, into an abelian group with O as its identity element.

We can also describe the group law algebraically and if you are not convinced that the above geometric description gives a group law, then you can carry out the tedious calculations required to verify that the algebraic formulas given below define a group structure on $E(L)$. Let $P_1 = (x_1, y_1), P_2 = (x_2, y_2) \in E(L)$. Then we define $-P_1 = (x_1, -y_1 - a_1x_1 - a_3)$. If $x_1 = x_2$ and $y_1 + y_2 + a_1x_2 + a_3 = 0$ then $P_1 + P_2 = O$ otherwise if $x_1 \neq x_2$ let

$$\lambda = \frac{y_2 - y_1}{x_2 - x_1} \tag{5.1}$$

$$\nu = \frac{y_1x_2 - y_2x_1}{x_2 - x_1}. \tag{5.2}$$

If $x_1 = x_2$ then let

$$\lambda = \frac{3x_1^2 + 2a_2x_1 + a_4 - a_1y_1}{2y_1 + a_1x_1 + a_3} \tag{5.3}$$

$$\nu = \frac{-x_1^3 + a_4x_1 + 2a_6 - a_3y_1}{2y_1 + a_1x_1 + a_3} \tag{5.4}$$

Then the line $y = \lambda x + \nu$ is a line through P_1, P_2 or is tangent to E at P_1 if $P_1 = P_2$. Then we have the following formula for $P_3 = (x_3, y_3) = P_1 + P_2$

$$x_3 = \lambda^2 + a_1\lambda - a_2 - x_1 - x_2 \tag{5.5}$$

$$y_3 = -(\lambda + a_1)x_3 - \nu - a_3 \tag{5.6}$$

Example 5.7. Let $K = \mathbb{Q}$, and let E/K be defined by

$$y^2 = x^3 - 6x^2 + 11x + 3 \tag{5.8}$$

then $\Delta_E = -34928, j_E = 6912/2183$.

Further $P_1 = (1,3), P_2 = (2,3)$ are $\mathbb{Q}$-rational points on E and $P_1 + P_2 = (3,-3)$ and $2P_1 = P_1 + P_1 = (17/4, -29/8)$, $3P_1 = 2P_1 + P_1 = (633/169, 6046/2197)$ and $4P_1 = (17889/13456, -3262625/1560896)$ etc.

When $L = \mathbb{C}$ the structure of $E(L)$ is completely understood using the theory of elliptic functions. The following theorem is a key result in the classical theory of elliptic functions.

Theorem 5.9. *Let $E/\mathbb{C}$ be an elliptic curve. Then there exists a lattice $L \subset \mathbb{C}$ (so $L \cong \mathbb{Z} \oplus \mathbb{Z}$ as abelian group) and a homomorphism of groups*

$$\mathbb{C}/L \cong E(\mathbb{C}) \tag{5.10}$$

which is given by $0 \mapsto O \in E(\mathbb{C})$ and $0 \neq z \mapsto (\mathfrak{p}_L(z), \mathfrak{p}'_L(z), 1)$ where $\mathfrak{p}_L$ and $\mathfrak{p}'_L$ are the Weierstrass elliptic functions with period lattice L.

Remark 5.11. Thus the group law on $E(L)$ gives us a way of generating more solutions from the given ones.

Remark 5.12. Fermat's method of infinite descent is also a variant of the group law on a suitable elliptic curve over $K = \mathbb{Q}$.

The following theorem due to Mordell, which was later generalized by André Weil to higher dimensional analogues of elliptic curves (see [47]) gives us a fundamental insight into the structure of the group $E(K)$. For a proof see [43]

Theorem 5.13 (Mordell-Weil)**.** *Let K be a number field and let L/K be a finite extension. Then for any elliptic curve E/K, the group $E(L)$ is a finitely generated abelian group, i.e.,*

$$E(L) \cong \mathbb{Z}^r \oplus \text{Finite group} \tag{5.14}$$

Remark 5.15. The number of copies of $\mathbb{Z}$ which occur in the above description is called the rank of $E(L)$. Birch and Swinnerton-Dyer have made a fascinating conjecture about the rank of $E(K)$ and the order of zero of a complex analytic function associated to E (see [1], [2] and [43]).

6. Elliptic curves over finite fields

In our discussion so far we have not specified the field K except in the examples. The formulae we have written down for the group law are valid over any field K. In this section we will assume that $K = \mathbb{F}_q$ a finite field with q elements and characteristic $p > 0$.

Suppose $E/\mathbb{F}_q$ is an elliptic curve over $\mathbb{F}_q$. Then it is easy to see that $E(\mathbb{F}_q)$ is a finite set as there are only finite number of possible values for each

coordinate of any $\mathbb{F}_q$-rational point. Hasse (see [23]) proved the following bound on the size of $E(\mathbb{F}_q)$ in terms of q. This estimate was later generalized by Weil in [48]. More precisely, Hasse proved the following:

Theorem 6.1 (Hasse-Weil Estimate)**.** *Let $E/\mathbb{F}_q$ be an elliptic curve over a finite field with q elements. Then we have*

$$|\#E(\mathbb{F}_q) - q - 1| \leq 2q^{1/2} \tag{6.2}$$

In other words if we write

$$\#E(\mathbb{F}_q) = q + 1 - a_q \tag{6.3}$$

then

$$|a_q| \leq 2q^{1/2}. \tag{6.4}$$

For a proof of the Hasse-Weil estimate, see [43].

This estimate was later generalized by Weil to arbitrary smooth projective curves and to abelian varieties and higher dimensional varieties. Weil's paper [48] is an excellent and elementary introduction to the subject. A more modern and elementary account can be found in [26].

Some of the Weil conjectures were proved by [18] and [31]. In 1974 Deligne proved the Weil conjectures (see [12]). For a reader with some basic background in algebraic geometry we recommend Katz's exposition of Deligne's proof (see [27]).

Remark 6.5. Elliptic curves over finite fields play an important role in primality testing and cryptography (see [28]).

7. Minimal Equations

Let E be an elliptic curve over $\mathbb{Q}$. We will say that the equation defining E is *minimal* at a prime $p|\Delta_E$ if the valuation $\nu_p(\Delta_E)$ is least amongst all possible choices of equations for E such that coefficients are p-adic integers. By formulas (3.18), we see that we can change the equation of E and we can get $\Delta' = u^{-12}\Delta$. Performing this change of coordinates as many times as possible we arrive at $\Delta_p(E) < 12$ and such that all the coefficients are p-adic integers. Thus we can conclude: if $v_p(a_i) \geq 0$ and $\nu_p(\Delta) < 12$ then this equation of E is minimal.

We note that $v_p(a_i) \geq 0$ and $\nu_p(\Delta_E) < 12$ is a sufficient condition for minimality of the equation.

Remark 7.1. Let E/K be an elliptic curve over a number field K and let $\mathcal{O}_K$ be the ring of integers and $\mathfrak{p} \subset \mathcal{O}_K$ a prime ideal. We can also formulate the notion of a minimal equation for E at $\mathfrak{p}$ exactly as above.

Example 7.2. Let $y^2 = x^3 + 16$, then this equation define an elliptic curve over $\mathbb{Q}$ and $\Delta_E = -2^{12}3^3$. The equation is not minimal at $p = 2$ and the transformation $x = 4x'$ $y = 8y' + 4$ transforms the equation of E to $(y')^2 + y' = (x')^3$. This equation has $\Delta = -27$, and is minimal at 2 and in fact at all primes.

It is not difficult to see using (3.18) that any equation E/K, where $K/\mathbb{Q}_p$ is a finite extension of $\mathbb{Q}_p$, has a minimal equation over K. The following proposition is easy to prove using the transformations (3.18) and the fact that every prime ideal in $\mathbb{Z}$ is principal.

Proposition 7.3. *Let $E/\mathbb{Q}$ be an elliptic curve over $\mathbb{Q}$. Then there exists an equation with integer coefficients for E which is minimal at all primes.*

From now on we will assume that all our elliptic curves are given by a minimal equation.

8. Reducing modulo primes

Let K be a number field, let $\mathcal{O}_K$ be the ring of integers of K and $\mathfrak{p}$ be a non-zero prime ideal in $\mathcal{O}_K$. Let E/K be an elliptic curve defined by an equation which is a *minimal equation* at $\mathfrak{p}$.

$$y^2 + a_1xy + a_3 = x^3 + a_2x^2 + a_4x + a_6 \tag{8.1}$$

We will assume for simplicity that $a_i \in \mathcal{O}_K$. Then we can reduce the equation defining E modulo $\mathfrak{p}$ and arrive a new equation which has coefficients in $k = \mathcal{O}_K/\mathfrak{p}$ which is a finite field:

$$y^2 + \bar{a}_1xy + \bar{a}_3 = x^3 + \bar{a}_2x^2 + \bar{a}_4x + \bar{a}_6 \tag{8.2}$$

where $\bar{a_i} = a_i \mod \mathfrak{p}$.

Observe that (8.2) represents an elliptic curve over $k = \mathcal{O}_K/\mathfrak{p}$ if and only if the discriminant $\bar{\Delta} = \Delta \mod \mathfrak{p}$ is not zero, i.e., provided that $\Delta \notin \mathfrak{p}$ or equivalently $\mathfrak{p} \not| \Delta$.

Thus this recipe of reducing the equation of an elliptic curve over a number field produces an elliptic curve over the finite field $\mathcal{O}_K/\mathfrak{p}$ provided $\mathfrak{p} \not| \Delta$. As Δ is divisible by only a finite number of primes we see that for all but finite number of primes in K we will get an elliptic curve over the corresponding finite field.

Definition 8.3. In the notations of the above paragraphs, we will say that $\mathfrak{p}$ is a prime of *good reduction if* $\mathfrak{p} \not| \Delta$.

Definition 8.4. If $\mathfrak{p}|\Delta$ then we will say that $\mathfrak{p}$ is a prime of *bad reduction* for E.

If $\mathfrak{p}$ is a prime of bad reduction then our recipe fails to produce an elliptic curve but produces a singular curve instead. By using our discussion of singular curves (see Section 4) we can further classify primes of bad reduction. We know from Section 4 that the reduction of E has either a node or a cusp.

Definition 8.5. In the notations as above, We will say that E has *semistable* or *multiplicative* reduction at $\mathfrak{p}$ if $\mathfrak{p}|\Delta$ and the reduction of E modulo $\mathfrak{p}$ has a node. If the two tangents at the node are defined over $\mathcal{O}_K/\mathfrak{p}$ then we say that E has *split multiplicative reduction* at $\mathfrak{p}$. If the two tangents at the node are not defined over $\mathcal{O}_K$ then we say that E has *non-split multiplicative reduction* at $\mathfrak{p}$. If the reduction of E at $\mathfrak{p}$ is a cusp then we say that E has *additive* or *unstable* reduction at $\mathfrak{p}$.

Proposition 8.6. *Let E/K be an elliptic curve defined by a minimal equation*

$$y^2 + a_1xy + a_3 = x^3 + a_2x^2 + a_4x + a_6 \tag{8.7}$$

with $a_i \in \mathcal{O}_K$. Let $\mathfrak{p}$ be a prime in $\mathcal{O}_K$.

1. *E has good reduction at $\mathfrak{p}$ if and only if $\Delta_E \not\equiv 0 \mod \mathfrak{p}$,*
2. *E has additive reduction modulo $\mathfrak{p}$ if and only if $\Delta_E \equiv c_4 \equiv 0 \mod \mathfrak{p}$,*
3. *E has semistable reduction modulo $\mathfrak{p}$ if and only if $\Delta_E \equiv 0 \mod \mathfrak{p}$ and $c_4 \not\equiv 0 \mod \mathfrak{p}$.*

Remark 8.8. Additive reduction is the worse kind of bad reduction while semistable reduction is not too bad. The following may help illustrate the subtle difference between these two types of bad reduction. Let E/K be an elliptic curve and let $\mathfrak{p}$ be a prime of bad reduction for E. Let K'/K be a finite extension, $\mathfrak{p}'$ be any prime lying over $\mathfrak{p}$ in K'. If E has semistable reduction at $\mathfrak{p}$ then, E thought of as an elliptic curve over K', continues to have semistable reduction at $\mathfrak{p}'$. However, if E has additive reduction at $\mathfrak{p}$, then there exists a finite extension K'/K such that for any prime $\mathfrak{p}'$ lying over $\mathfrak{p}$, E has good or semistable reduction at $\mathfrak{p}'$. In other words, additive reduction may disappear or become semistable over a suitable finite extension, while semistable reduction persists. The following examples illustrate this point further.

Example 8.9. Let $K = \mathbb{Q}$, let E be defined by $y^2 = x^3 + x^2 + 17$. Then E has multiplicative reduction modulo 17.

Example 8.10. Let $K = \mathbb{Q}$ and let E be defined by $y^2 = x^3 + 17$. Then E has additive reduction modulo 17. Let $K' = \mathbb{Q}(17^{1/6})$, and then we can write the equation of E over K' as

$$(y/(17^{1/6})^3)^2 = (x/(17^{1/6})^2)^3 + 1 \tag{8.11}$$

Writing $y' = y/17^{1/2}, x' = x/17^{1/3}$ we get $y'^2 = x'^3 + 1$ and this new equation has good reduction modulo the unique prime lying over 17 in $\mathbb{Q}(17^{1/6})$.

9. The conductor of an elliptic curve

Let K be a number field. Let E/K be an elliptic curve. We can define an ideal, $N_E \subset \mathcal{O}_K$ of the ring of integers of K,called the conductor ideal or more simply the conductor of E. This ideal is defined as

$$N_E = \prod_{\mathfrak{p}|\Delta} \mathfrak{p}^{f_{\mathfrak{p}}(E/K)} \tag{9.1}$$

where the exponents $f_{\mathfrak{p}}(E/K)$ of $\mathfrak{p}$ are defined as follows.

$$f_{\mathfrak{p}}(E/K) = \begin{cases} 0 & \text{if } E \text{ has good reduction at } \mathfrak{p}, \\ 1 & \text{if } E \text{ has semistable reduction at } \mathfrak{p}, \\ 2 & \text{if } \mathfrak{p} \not| 6 \text{ and } E \\ & \qquad \text{has additive reduction at } \mathfrak{p}, \\ 2 + \delta_{\mathfrak{p}}(E/K) & \text{if } \mathfrak{p}|6. \end{cases} \tag{9.2}$$

where the exact definition of $\delta_{\mathfrak{p}}(E/K)$ is irrelevant for our purposes. No exact formulas are known for $\delta_{\mathfrak{p}}(E/K)$, but Tate's algorithm (see [45]) gives a way of computing $f_{\mathfrak{p}}$ for all primes including those dividing 6. It is known, for instance, that

$$f_2(E/\mathbb{Q}) \leq 8 \tag{9.3}$$

$$f_3(E/\mathbb{Q}) \leq 5 \tag{9.4}$$

Observe that the conductor ideal is divisible by only those primes which divide Δ_E, and that it captures finer reduction information of E/K. The conductor is one of the fundamental arithmetical invariants of an elliptic curve.

Remark 9.5. Tate's algorithm [45], and other algorithms like [30] have now been implemented in many software packages (notably in the software package PARI-GP which is available on Internet; this package also contains an extensive elliptic curve computation package which computes many numerical invariants of elliptic curves).

Example 9.6. Let $y^2 = x^3 + 10x + 11$. Then $\Delta_E = -2^4 \cdot 13^2 \cdot 43$ and $N_E = 2^4 \cdot 13 \cdot 43$ and the equation is minimal.

10. Action of the Galois group

Let K be a number field or a finite extension of $\mathbb{Q}_p$ or a finite field. Let $\bar{K}$ be an algebraic closure of K. We write $G_K = \mathrm{Gal}(\bar{K}/K)$ for the Galois group of $\bar{K}/K$. Let E/K be an elliptic curve and suppose $P = (x_0, y_0) \in E(\bar{K})$. Thus P satisfies

$$y_0^2 + a_1 x_0 y_0 + a_3 = x_0^3 + a_2 x_0^2 + a_4 x_0 + a_6. \tag{10.1}$$

Suppose $\sigma \in G_K$. We apply σ to the above equation and get

$$\sigma(y_0)^2 + a_1 \sigma(x_0)\sigma(y_0) + a_3 = \sigma(x_0)^3 + a_2 \sigma(x_0)^2 + a_4 \sigma(x_0) + a_6. \tag{10.2}$$

where we have used the fact that $\sigma(a_i) = a_i$ for any $a_i \in K$. Thus we see from the above equation that $P^\sigma = (\sigma(x_0), \sigma(y_0)) \in E(\bar{K})$ is also a point on the elliptic curve. In other words, G_K operates on $E(\bar{K})$.

We now study this action in some more detail. Let $n \geq 2$ be any integer and let

$$E[n] = \{P \in E(\bar{K}) | nP = O\} \tag{10.3}$$

Then it is easy to see that $E[n] \subset E(\bar{K})$ is a subgroup. We call $E[n]$ the group of *n-torsion points* on the curve E or simply the group of points of order dividing n.

The structure of $E[n]$ as an abelian group is completely understood.

Theorem 10.4. *Let K be a field and let E be an elliptic curve over K. If the characteristic of K does not divide n then one has*

$$E[n] \cong \mathbb{Z}/n \oplus \mathbb{Z}/n. \tag{10.5}$$

In particular for for a number field K, and when $n = p$ a prime, we see that $E[p]$ is a two dimensional vector space over $\mathbb{F}_p$.

When the characteristic of K divides n the result is slightly different but as we will not need it here we do not recall it and the reader is referred to [43] for more details.

Moreover, if $P \in E[n]$ and $\sigma \in G_K$ then

$$\sigma(nP) = n\sigma(P) = O \tag{10.6}$$

So that G_K acts on the $\mathbb{Z}/n$-module $E[n]$.

Example 10.7. Let $K = \mathbb{Q}$ and let E be defined by $y^2 = x^3 - x$. Then $E[2] = \{O, (0,0), (0,1), (0,-1)\}$ and so $G_{\mathbb{Q}}$ operates trivially on $E[2]$. For a more interesting example see Remark 11.20.

11. Tate Elliptic curves

Let $K = \mathbb{Q}_p$ be the field of p-adic numbers, where p is any prime, let $|\ |_p$ denote the p-adic absolute value, and let $\nu_p : \mathbb{Q}_p \to \mathbb{Z}$ be the normalized p-adic valuation, normalized so that $\nu_p(p) = 1$. As $\mathbb{Q} \subset \mathbb{Q}_p$, we can think of any elliptic curve over $\mathbb{Q}$ as an elliptic curve over $\mathbb{Q}_p$.

Tate discovered elliptic curves over $\mathbb{Q}_p$ with remarkable properties. These curves are called Tate elliptic curves (see [44]). These curves are sort of "universal models" for elliptic curves over $\mathbb{Q}_p$ with split multiplicative reduction modulo p. Tate curves are given by an explicit equation.

Fix $q \in \mathbb{Q}_p$ such that $|q|_p < 1$, so q is a p-adic integer divisible by p. Let E_q be defined by

$$y^2 + xy = x^3 + a_4(q)x + a_6(q) \tag{11.1}$$

where

$$a_4(q) = -5\sum_{n=1}^{\infty} \frac{n^3 q^n}{1-q^n} \tag{11.2}$$

and

$$a_6(q) = -\frac{1}{12}\sum_{n=1}^{\infty} \frac{(7n^5+5n^3)q^n}{(1-q^n)} \tag{11.3}$$

Then the discriminant $\Delta_{E_q} = \Delta(q)$ is the famous Ramanujan function

$$\Delta(q) = q\prod_{n=1}^{\infty}(1-q^n)^{24} \tag{11.4}$$

and the j-invariant is

$$j(q) = \frac{1}{q} + 744 + 196884q + \cdots \tag{11.5}$$

Remark 11.6. It is clear from the definition of discriminant $\Delta(q)$ that

$$\nu_p(\Delta(q)) = \nu_p(q) \tag{11.7}$$

$$\nu_p(j(q)) = -\nu_p(q) \tag{11.8}$$

Further for $|q|_p < 1$, the series $a_4(q)$ and $a_6(q)$ converge and (11.1) has points in $\mathbb{Q}_p((u))$ given by convergent power series in u

$$x(q,u) = \sum_{n\in\mathbb{Z}} \frac{q^n u}{(1-q^n u)^2} - 2\sum_{n=1}^{\infty}\frac{nq^n}{1-q^n} \tag{11.9}$$

$$y(q,u) = \sum_{n\in\mathbb{Z}} \frac{q^{2n}u^2}{(1-q^n u)^3} + \sum_{n=1}^{\infty}\frac{nq^n}{1-q^n} \tag{11.10}$$

Then for any $|q|_p < 1$ and $u \in \bar{\mathbb{Q}}_p^*$ these power series converge and their values give a point $(x, y) \in E_q(\bar{\mathbb{Q}}_p)$. Moreover if we reduce the equation of E_q modulo p then we get

$$y^2 + xy = x^3 \tag{11.11}$$

and so one checks easily from this that E_q has split multiplicative reduction modulo p.

Observe that $x(q, u)$ and $y(q, u)$ as a function of u has the property that

$$x(q, u) = x(q, qu) \tag{11.12}$$

$$y(q, u) = y(q, qu). \tag{11.13}$$

Thus these functions are periodic with respect to the multiplicative group

$$q^{\mathbb{Z}} = \{q^m | m \in \mathbb{Z}\}. \tag{11.14}$$

Tate further showed that if we fix q with $|q|_p < 1$ then the mapping

$$\bar{\mathbb{Q}}_p^* \to E_q(\bar{\mathbb{Q}}_p) \tag{11.15}$$

given by $u \mapsto (x(q, u), y(q, u))$ is a surjective homomorphism of groups with kernel $q^{\mathbb{Z}}$ and is compatible with the action of Galois group $\mathrm{Gal}(\bar{\mathbb{Q}}_p/\mathbb{Q}_p)$ on both the sides. Thus one gets

$$\mathbb{Q}_p^*/q^{\mathbb{Z}} \cong E_q(\mathbb{Q}_p) \tag{11.16}$$

The crucial property of Tate curves is described in the following theorem.

Theorem 11.17 (Tate uniformization)**.** *Let $E/\mathbb{Q}$ be an elliptic curve and suppose that $p|\Delta_E$. Assume that E has split multiplicative reduction at p. Then there exists a $q \in \mathbb{Q}_p$ such that $E \cong E_q$ over $\mathbb{Q}_p$.*

Tate proved a result which is more general than the above statement. In particular, Tate's result is valid for curves over number fields and finite extensions of $\mathbb{Q}_p$ (see [43]). Moreover, one also get a similar assertion for non-split multiplicative reduction. But we will not need the general assertion here and we refer the reader to [43] for more details.

Example 11.18. Let $K = \mathbb{Q}_p$ and let E_q be a Tate curve over K. We can use Tate's theorem to describe $E[n]$ explicitly. The $u \in \bar{\mathbb{Q}}_p^*$ has the property that

$$nP = n(x(q, u), y(q, u)) = O \tag{11.19}$$

if and only if $u^n \in q^{\mathbb{Z}}$, i.e., $nP = O$ if and only if $u^n = q^m$ for some $m \in \mathbb{Z}$ and so $u = \zeta q^{m/n}$ for some n^{th} root of unity ζ. Thus we obtain an isomorphism

$$E[n] \cong \left\{\zeta u^{m/n} \middle| \zeta^n = 1, 0 \leq m \leq n\right\}. \tag{11.20}$$

12. Some Galois Theory

In this section we will recall a few facts about the structure of the Galois Groups of number fields or local fields. For proofs or details see [40].

Fix prime p. We would like to recall a few facts about the structure of the Galois group $G_{\mathbb{Q}} = \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$. We will do this by studying the structure of $G_p = \mathrm{Gal}(\bar{\mathbb{Q}}_p/\mathbb{Q}_p)$. Let $\mathfrak{p}|p$ be any valuation lying over p in $\bar{\mathbb{Q}}$. The Galois group $G_{\mathbb{Q}}$ acts on the set of such $\mathfrak{p}$ transitively. Fix one such valuation $\mathfrak{p}|p$.

Definition 12.1. The decomposition group $D(\mathfrak{p}, p) \subset G_{\mathbb{Q}}$ at $(\mathfrak{p}, p)$ is defined to be the set of all $\sigma \in G_{\bar{\mathbb{Q}}}$ such that $\sigma(\mathfrak{p}) = \mathfrak{p}$.

In other words, $D(\mathfrak{p}, p)$ is the stabilizer of $\mathfrak{p}$. The decomposition group depends on $\mathfrak{p}$ and p. If we replace $\mathfrak{p}$ by another $\mathfrak{p}'|p$ then the decomposition group $D(\mathfrak{p}', p)$ is a conjugate of $D(\mathfrak{p}, p)$. Hence we will often suppress the dependence of $\mathfrak{p}$ and often call $D(\mathfrak{p}, p)$ the decomposition group at p.

The decomposition group encodes a lot of information about p, it contains interesting subgroups which encode information about the ramification of p in any extension. The following result identifies the decomposition group little more explicitly.

Proposition 12.2. *For every prime $\mathfrak{p}|p$ in $\bar{\mathbb{Q}}$, we have an isomorphism*

$$\mathrm{Gal}(\bar{\mathbb{Q}}_p/\mathbb{Q}_p) \cong D(\mathfrak{p}, p) \hookrightarrow \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}). \tag{12.3}$$

So for every prime $\mathfrak{p}|p$ in $\bar{\mathbb{Q}}$ we get an embedding of $\mathrm{Gal}(\bar{\mathbb{Q}}_p/\mathbb{Q}_p) \hookrightarrow \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$.

Thus we reduce to the of study the Galois group of $\bar{\mathbb{Q}}_p/\mathbb{Q}_p$. From now on we study these groups.

We have a natural surjection

$$G_p \to \mathrm{Gal}(\bar{\mathbb{F}}_p/\mathbb{F}_p) \tag{12.4}$$

which is given by $\sigma \mapsto \sigma \mod p$.

Definition 12.5. The kernel $G_p \to \mathrm{Gal}(\bar{\mathbb{F}}_p/\mathbb{F}_p)$ of

$$I_p = \{\sigma \equiv 1 \mod p\} \tag{12.6}$$

is called the inertia subgroup at p.

Thus we have an isomorphism

$$G_p/I_p \cong \mathrm{Gal}(\bar{\mathbb{F}}_p/\mathbb{F}_p), \tag{12.7}$$

given by $\sigma \mapsto \sigma \mod p$.

The $\mathrm{Gal}(\bar{\mathbb{F}}_p/\mathbb{F}_p)$ has natural element $\mathrm{Frob}_p \in \mathrm{Gal}(\bar{\mathbb{F}}_p/\mathbb{F})$ which is given by raising to p^{th}-powers:

$$\mathrm{Frob}_p(x) = x^p \tag{12.8}$$

for all $x \in \bar{\mathbb{F}}_p$. We will call Frob_p the Frobenius (morphism) at p.

A proof of the following proposition can be found in [40].

Proposition 12.9. *Let L/K be any Galois extension of number fields. Let $\mathfrak{p}'\mathfrak{p}$ be a prime of L lying over a prime $\mathfrak{p}$ of K. Then L/K is unramified at $\mathfrak{p}$ if and only if the inertia subgroup $I_\mathfrak{p}$ at $\mathfrak{p}$ is trivial.*

Definition 12.10. Let $L/\mathbb{Q}$ be an arbitrary Galois extension. Assume that $L/\mathbb{Q}$ is unramified outside a finite set of primes. Let p be a prime at which L is unramified. Let $\mathfrak{p}$ be a prime lying over p in L. Then a Frobenius element at p is a conjugacy class of any element $\sigma \in \mathrm{Gal}(L/\mathbb{Q})$ such that σ is in $D(\mathfrak{p}, p)$ and its image in $\mathrm{Gal}(\bar{\mathbb{F}}_p/\mathbb{F}_p)$ under the isomorphism (12.7) is $\mathrm{Frob}_p \in \mathrm{Gal}(\bar{\mathbb{F}}_p/\mathbb{F}_p)$.

We will continue to use the notation Frob_p to denote Frobenius element at p, keep in mind that it is really a conjugacy class of elements which depends only on p and not on the choice of a prime lying over p in L.

The following variant of the Chebotarev density theorem (see [6] or [38]) will be used without proof.

Theorem 12.11. *Let $L/\mathbb{Q}$ be an arbitrary Galois extension which is unramified outside a finite set of primes. Then the the set of Frobenius elements of primes which are unramified in $L/\mathbb{Q}$ is dense in* $\mathrm{Gal}(L/\mathbb{Q})$.

The inertia subgroup I_p has finer structure as well. We can define a filtration

$$I_i = \left\{\sigma | \sigma \equiv 1 \mod p^{i+1}\right\} \tag{12.12}$$

for all $i \geq 0$, with $I_0 = I_p$. There is a natural surjection from

$$I_p \to \bar{\mathbb{F}}_p^* \tag{12.13}$$

given by the action of the inertia group on the roots of unity in $\bar{\mathbb{Q}}_p$ and it is standard that

$$I_1 = \ker(I_p \to \bar{\mathbb{F}}_p^*). \tag{12.14}$$

The quotient group

$$I_t = I_p/I_1 \tag{12.15}$$

is often called the *tame quotient of I_p and I_1 is called the* wild inertia subgroup.

13. Galois Representations

Let E/K be an elliptic curve and assume that K is a field of characteristic zero. Let $n \geq 2$ be an integer. Then as we have seen that G_K operates on $E[n] \cong \mathbb{Z}/n \times \mathbb{Z}/n$. This action is compatible with the group structure and so for each $\sigma \in G_K$, the mapping $P \mapsto \sigma(P)$ is an automorphism of the $\mathbb{Z}/n$-module $E[n]$. Thus we get a homomorphism

$$\rho : G_K \to \mathrm{Aut}(E[n]) = \mathrm{GL}_2(\mathbb{Z}/n) \tag{13.1}$$

where

$$\mathrm{GL}_2(\mathbb{Z}/n) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} | a, b, c, d \in \mathbb{Z}/n \text{ and } ad - bc \in (\mathbb{Z}/n)^* \right\} \tag{13.2}$$

is the group of invertible matrices with coefficients in $\mathbb{Z}/n$.

Elliptic curves are sources of such homomorphisms but there are other sources of such homomorphisms and the confluence of two such sources is essentially the content of Serre's conjecture.

Definition 13.3. Let R be any ring, and let K be either a number field or a finite extension of $\mathbb{Q}_\ell$ for a prime ℓ. A Galois representation of G_K is a continuous homomorphism

$$\rho : G_K \to \mathrm{GL}_n(R) \tag{13.4}$$

where $\mathrm{GL}_n(R)$ denotes the group of $n \times n$ invertible matrices with entries in R. If R has a natural topology then we give $\mathrm{GL}_n(R)$ the topology induced on it as an open subset of R^n, and otherwise we give R the discrete topology.

Example 13.5. We will be interested in the situation when R is either the field of p-adic numbers for some p or is $\mathbb{Z}/n$ for some integer n. In the first case we give $\mathrm{GL}_n(\mathbb{Q}_p)$ the topology induced from $\mathbb{Q}_p^n$ and then continuity condition is with respect to this topology.

Example 13.6. When $R = \mathbb{Z}/n$, then $\mathrm{GL}_n(R)$ is a finite group and we give it the discrete topology. Continuity of ρ then simply means that $\ker(\rho)$ is an open subgroup of G_K.

Remark 13.7. It is often convenient to think of $\mathrm{GL}_n(R)$ as the group invertible R-linear maps from an R-module $V = R^n$ to itself. Any representation $\rho : G_K \to \mathrm{GL}_n(R)$ thus gives rise to an R-linear action of G_K on R^n, i.e., for each $\sigma \in G_K$ we have an invertible R-linear mapping $R^n \to R^n$ which is continuous and which satisfies obvious conditions and conversely any such action gives rise to a representation ρ.

Example 13.8. Let $\mu_n = \left\{ z \in \bar{K} | z^n = 1 \right\}$ be the group of n^{th} roots of unity in $\bar{K}$. Then G_K acts on it. We have an isomorphism of abelian groups $\mu_n \cong \mathbb{Z}/n$. Thus this action gives rise to a homomorphism

$$G_K \to \mathrm{Aut}(\mu_n) = \mathrm{GL}_1(\mathbb{Z}/n) = (\mathbb{Z}/n)^*. \tag{13.9}$$

Example 13.10. We will need the following special case of the above example. Suppose $n = \ell$ for a prime ℓ and

$$\chi_\ell : G_K \to \mathrm{Aut}(\mu_\ell) = (\mathbb{Z}/\ell)^*. \tag{13.11}$$

This representation is called the *cyclotomic character* at ℓ.

Example 13.12. Let ℓ be a prime and let $K = \mathbb{Q}$. Let us examine $\chi_\ell : G_{\mathbb{Q}} \to (\mathbb{Z}/\ell)^*$ in a little more detail. Let ζ be any ℓ^{th}-root of unity; suppose that c is complex conjugation $c : \bar{\mathbb{Q}} \to \bar{\mathbb{Q}}$. Then we have $\chi(c)(\zeta) = \bar{\zeta} = \zeta^{-1}$ and in particular, we see that $\chi_\ell(c) = -1 \in (\mathbb{Z}/\ell)^*$.

Definition 13.13. We say that a representation $\rho : G_K \to \mathrm{GL}_n(R)$ is reducible if there exists a proper R-submodule $0 \neq W \subset V = R^n$ such that the action of ρ on V maps W into itself, i.e., for all $\sigma \in G_K$, we have $\sigma(W) \subset W$.

Definition 13.14. If $\rho : G_K \to \mathrm{GL}_n(R)$ is not reducible, then we say that ρ is an irreducible representation of G_K.

Example 13.15. Let E_q be a Tate curve over $\mathbb{Q}_p$. We had observed in Example 11.18 that $E_q[n] = \left\{ \zeta q^{m/n} \middle| \zeta^n = 1, 0 \leq m \leq n \right\}$. We claim that the representation of $G_p = \mathrm{Gal}(\bar{\mathbb{Q}}_p/\mathbb{Q}_p)$ is reducible. To see this we define a surjective homomorphism $E_q[n] \to \mathbb{Z}/n$ which is given by $\zeta q^{m/n} \mapsto m \mod n$. This homomorphism is clearly surjective and if we give the trivial action of G_p on $\mathbb{Z}/n$, then this homomorphism is also compatible with the action of G_p. Further we can also identify the kernel $\ker(E_q[n] \to \mathbb{Z}/n)$ with μ_n. Thus we have an exact sequence of abelian groups

$$0 \to \mu_n \to E_q[n] \to \mathbb{Z}/n \to 0 \tag{13.16}$$

and each map in the sequence is compatible with the action of G_p. Thus the representation of G_p on $E_q[n]$ is reducible as μ_n is a G_p-stable subspace of $E_q[n]$.

Example 13.17. Let E/K be an elliptic curve and let ℓ be a prime and $m \geq 1$ be any integer. Then we know that G_K operates on the group $E[\ell^m] \cong \mathbb{Z}/\ell^m \oplus \mathbb{Z}/\ell^m$. Thus we get a homomorphism $\rho_{\ell^m} : G_K \to \mathrm{Aut}(E[\ell^m]) \cong \mathrm{GL}_2(\mathbb{Z}/\ell^m)$.

It is not difficult to see that the composite map $G_K \to \mathrm{GL}_2(\mathbb{Z}/\ell^m) \to \mathrm{GL}_2(\mathbb{Z}/\ell^{m-1})$ is $\rho_{\ell^{m-1}}$. Thus we can put these representations together get a homomorphism $\rho : G_K \to \mathrm{GL}_2(\mathbb{Z}_\ell)$, and as $\mathrm{GL}_2(\mathbb{Z}_\ell) \hookrightarrow \mathrm{GL}_2(\mathbb{Q}_\ell)$ then further composition gives a representation $G_K \to \mathrm{GL}_2(\mathbb{Q}_\ell)$.

Proposition 13.18. *Let R be a field of characteristic zero and let $\rho_1, \rho_2 : G \to GL_n(R)$ be two continuous, finite dimensional irreducible representations of a topological group G. Suppose that the traces* $\mathrm{Trace}(\rho_1(g)) = \mathrm{Trace}(\rho_2(g))$ *for dense subset of $g \in G$. Then ρ_1 and ρ_2 are isomorphic representations.*

Proof. See [40] or [10]. □

14. Fine structure of Galois Representations

We want to use the structure of the Galois group $\mathrm{Gal}(\bar{\mathbb{Q}}_p/\mathbb{Q}_p)$ described in Section 12 to probe the structure of Galois representations.

Definition 14.1. Let $\rho : G_K \to \mathrm{GL}_n(R)$ be a continuous representation. Let $\mathfrak{p} \subset \mathcal{O}_K$ be a nonzero prime ideal in K. Then we say ρ is unramified at $\mathfrak{p}$ if the image, $\rho(I_\mathfrak{p})$, of the inertia subgroup, $I_\mathfrak{p}$ at $\mathfrak{p}$, under ρ is trivial.

Remark 14.2. The definition given above is independent of the choice of a prime lying over $\mathfrak{p}$ as the decomposition groups of all primes lying over $\mathfrak{p}$ in $\bar{K}$ are conjugate and so are the inertia subgroups.

Example 14.3. Let $\rho : G_K \to \mathrm{GL}_n(\mathbb{Z}/m)$ be a continuous representation. Let $H = \ker(\rho)$ be the kernel of ρ. Continuity of ρ shows that H is an open subgroup. Further as $\mathrm{GL}_n(\mathbb{Z}/m)$ is a finite group, we see that the image of ρ is a finite group as well. Let $K_\rho = \bar{K}^H$ be the fixed field of H. Galois Theory provides us an isomorphism:

$$\mathrm{image}(\rho) = \mathrm{Gal}(K_\rho/K) \tag{14.4}$$

Then $\rho : G_K \to \mathrm{GL}_m(\mathbb{Z}/n)$ is unramified at $\mathfrak{p}$ if and only if the extension K_ρ/K is unramified at $\mathfrak{p}$.

Remark 14.5. Let $\rho : G_\mathbb{Q} \to \mathrm{GL}_n(R)$ be any continuous Galois representation. Suppose ρ is unramified at p. Then $\rho(\mathrm{Frob}_p)$, where $frob_p$ is a Frobenius element at p, is conjugacy class of elements of $\mathrm{GL}_n(R)$. The characteristic polynomial $\det(1 - X\rho(\mathrm{Frob}_p))$ is well defined as it depends only on the conjugacy class of $\rho(\mathrm{Frob}_p)$. This characteristic polynomial plays a fundamental role in studying a Galois representations.

Example 14.6. Let $\chi_\ell : G_\mathbb{Q} \to (\mathbb{Z}/\ell)^*$ be the cyclotomic character. Then χ_ℓ is unramified at all primes $p \neq \ell$. We can calculate the Frobenius elements explicitly in this case. For all $p \neq \ell$, $\chi_\ell(\mathrm{Frob}_p) = p \mod \ell$ and hence $\det(1 - X\chi(\mathrm{Frob}_p)) = (1 - pX)$.

15. Modular forms

Let

$$\mathfrak{H} = \{z \in \mathbb{C} | \mathrm{Re}(z) > 0\} \tag{15.1}$$

be the upper half plane. Let

$$\mathrm{GL}_2^+(\mathbb{R}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} | a, b, c, d \in \mathbb{R}, ad - bc > 0 \right\} \tag{15.2}$$

We write

$$\mathrm{GL}_2^+(\mathbb{Q}) = \mathrm{GL}_2^+(\mathbb{R}) \cap \mathrm{GL}_2(\mathbb{Q}) \tag{15.3}$$

We let $\mathrm{GL}_2^+(\mathbb{R})$ act on the topological space $\mathfrak{H}$ by the formula

$$gz = \frac{az + b}{cz + d} \tag{15.4}$$

for any $g \in \mathrm{GL}_2^+(\mathbb{R})$ and any $z \in \mathfrak{H}$.

In this and the subsequent sections we will be interested in subgroups of finite index in $\mathrm{SL}_2(\mathbb{Z})$. We fix a subgroup $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ of finite index.

Definition 15.5. A modular form of weight k on Γ is a holomorphic function $f : \mathfrak{H} \to \mathbb{C}$ such that

(1) $f(gz) = (cz + d)^k f(z)$ for all $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$

(2) for every $g \in \mathrm{GL}_2^+(\mathbb{Q})$ the function $\det(g)^{k/2}(cz + d)^{-k} f(gz)$, where $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ has a Fourier expansion of the form

$$\det(g)^{-k}(cz + d)^{-k} f(gz) = \sum_{n=0}^{\infty} a_n e^{2\pi i n z/N} \tag{15.6}$$

for some integer $N \geq 1$.

Definition 15.7. We say that a modular form of weight k on Γ is a cusp form if in addition to the above two conditions we have

$$\det(g)^{-k}(cz + d)^{-k} f(gz) = \sum_{n=1}^{\infty} a_n e^{2\pi i n z/N} \tag{15.8}$$

It is easy to verify that a linear combination of modular forms on Γ of weight k is again a modular form on Γ of weight k and product of two modular forms of weights k and m is a modular form of weight $k + m$ on Γ. Thus the set of modular forms of fixed weight on Γ form a complex vector space and the set of all cusp forms is a subspace of the space modular forms.

Example 15.9. Let $k > 2$ be an even integer. For any $z \in \mathfrak{H}$ consider the series

$$G_k(z) = \sum_{(m,n)\neq(0,0)}{}' \frac{1}{(mz+n)^k} \tag{15.10}$$

and the sum is over all integers (m, n) which are not simultaneously zero. Then it is easy to see that the series is absolutely convergent for all values of $z \in \mathfrak{H}$ and one has $G_k(gz) = (cz + d)^k G_k(z)$ for all $g \in \mathrm{SL}_2(\mathbb{Z})$. Its Fourier expansion is given by

$$G_k(z) = 2\zeta(k)\left(1 - \frac{2k}{B_k}\sum_{n=1}^{\infty}\sigma_{k-1}(n)q^n\right) \tag{15.11}$$

and where $q = e^{2\pi iz}$. Thus $G_k(z)$ is a modular form of weight k on $\mathrm{SL}_2(\mathbb{Z})$. But $G_k(z)$ is not a cusp form.

Example 15.12. The Ramanujan $\Delta(q)$ function defined by

$$\Delta(q) = q\prod_{n=1}^{\infty}(1-q^n)^{24} = \sum_{n} = 1^{\infty}\tau(n)q^n.$$

where $q = e^{2\pi iz}$ is a cusp form of weight 1 on $\mathrm{SL}_2(\mathbb{Z})$.

Definition 15.13. Let Γ be a subgroup of finite index in $\mathrm{SL}_2(\mathbb{Z})$. Let $k \geq 1$ be an integer. Let $M_k(\Gamma)$ (resp. $S_k(\Gamma)$) be the space of modular forms of weight k (resp. cusp forms) on weight k on Γ.

Observe that $S_k(\Gamma) \subset M_k(\Gamma)$.

From now on we will be interested in the following kinds of subgroups of $\mathrm{SL}_2(\mathbb{Z})$.

Definition 15.14. Let $N \geq 1$ be any integer. Define $\Gamma_1(N)$ as follows

$$\Gamma_1(N) = \left\{g = \begin{pmatrix}1 & * \\ 0 & 1\end{pmatrix} \mod N \middle| g \in \mathrm{SL}_2(\mathbb{Z})\right\} \tag{15.15}$$

and

$$\Gamma_0(N) = \left\{g = \begin{pmatrix}* & * \\ 0 & *\end{pmatrix} \mod N \middle| g \in \mathrm{SL}_2(\mathbb{Z})\right\} \tag{15.16}$$

I leave it as an exercise to check that these two subgroups are of finite index in $\mathrm{SL}_2(\mathbb{Z})$. Moreover one has a homomorphism

$$\Gamma_0(N) \to (\mathbb{Z}/N)^* \tag{15.17}$$

given by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto d \mod N. \tag{15.18}$$

This defines a surjective homomorphism of groups and the kernel of this homomorphism is precisely the subgroup $\Gamma_1(N)$ and hence we see that $\Gamma_1(N) \subset \Gamma_0(N)$ is a normal subgroup.

We will be interested in studying modular forms on these two groups. It is standard (see [42]) that $M_k(\Gamma_1(N))$ and $M_k(\Gamma_0(N))$ are finite dimensional $\mathbb{C}$-vector spaces. In particular the spaces of cusp forms on these groups are finite dimensional as well.

Definition 15.19. A modular form (resp. cusp form) of level N, weight k is a form $f \in M_k(\Gamma_1(N))$ (resp. $f \in S_k(\Gamma_1(N))$).

Let $\chi : (\mathbb{Z}/N) \to \mathbb{C}^*$ be any homomorphism (i.e., a Dirichlet character). For $d|N$ we set $\chi(d) = 0$ and extend this function to $\mathbb{Z}/N$.

Definition 15.20. A modular form $f \in M_k(\Gamma_1(N))$ is said to be a modular form of level N, weight k and *nebentype* χ if f satisfies the following:

$$f(gz) = \chi(d)(cz + d)^k f(z) \tag{15.21}$$

and we write $M_k(\Gamma_0(N), \chi) \subset M_k(\Gamma_1(N))$ for the space of such forms on $\Gamma_1(N)$. We also define $S_k(\Gamma_0(N), \chi)$ in the obvious way.

One has the following decomposition of complex vector spaces

$$M_k(\Gamma_1(N)) = \oplus_{\chi \mod N} M_k(\Gamma_0(N), \chi), \tag{15.22}$$

and

$$S_k(\Gamma_1(N)) = \oplus_{\chi \mod N} S_k(\Gamma_0(N), \chi). \tag{15.23}$$

for space of cusp forms.

Definition 15.24. A modular form (resp. cusp form) of level N, weight k and nebentype χ is a form $f \in M_k(\Gamma_0(N), \chi)$ (resp. $f \in S_k(\Gamma_0(N), \chi)$).

Example 15.25. Let $N = 11$. Then the function $f(z) = (\Delta(q)\Delta(11q)^{1/2} = q\prod_{n=1}^{\infty}((1-q^n)(1-q^{11n})^2$ is a cusp form of level 11, weight 2 and nebentype $\chi = 1$.

Remark 15.26. Let $f(z) \in M_k(\Gamma_1(N))$ be a modular form of weight k and level N. Then as

$$T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \Gamma_1(N) \tag{15.27}$$

we see that $f(Tz) = f(z+1) = f(z)$ so $f(z)$ is a complex holomorphic function which is periodic with period 1 and so its Fourier expansion looks like

$$f(z) = \sum_{n=1}^{\infty} a_n e^{2\pi i n z} \tag{15.28}$$

Definition 15.29. We will call the coefficients a_i the Fourier coefficients of f and write $K_f = \mathbb{Q}(a_0, a_1, \cdots)$ for the field generated by the Fourier coefficients of f. In general K_f is not even an algebraic extension of $\mathbb{Q}$. But under interesting circumstances it is a finite extension of $\mathbb{Q}$. In any case we will refer to K_f as the field of Fourier coefficients of f.

16. Hecke operators

We will restrict attention to the two special subgroups introduced earlier. For a general and comprehensive account of the theory see [42], [25], [29]. From now on we will concentrate on the space of cusp forms on $\Gamma_1(N)$ and keep a track of the nebentype as we go along.

The spaces of modular forms on $\Gamma_1(N)$ come equipped with a commutative family of operators, which were introduced by Hecke and are named after him. We will not recall the definition of Hecke operators but only recall the effect of Hecke operators on the Fourier coefficients of any modular form. For every $n \geq 1$ we have a $\mathbb{C}$-linear mapping

$$T_n : M_k(\Gamma_1(N)) \to M_k(\Gamma_1(N)) \tag{16.1}$$

which takes the subspace of cusp forms to cusp forms and forms withe Nebentype to forms with the same nebentype. We call T_n the n^{th} Hecke operators. First property of these operators is

$$T_{nm} = T_n T_m \qquad \text{for } (m,n) = 1, \tag{16.2}$$

Thus it is sufficient to define these operators when $n = \ell^m$ for any prime ℓ and $m \geq 1$. So fix a prime ℓ. Let $f \in S_k(\Gamma_0(N), \chi)$ be a form level N, weight k and nebentype χ.

$$T_\ell(f) = \sum_{n=1}^{\infty} a_{\ell n} q^n + \chi(\ell)\ell^{k-1} \sum_{n=1}^{\infty} a_n q^{\ell n} \tag{16.3}$$

recall that when $\ell | N$ we have set $\chi(\ell) = 0$.

We define T_{ℓ^m} recursively using the above definition for $m = 1$.

Definition 16.4. A Hecke eigenform $f \in S_k(\Gamma_0(N), \chi)$ is a common eigen function of all the Hecke operators T_ℓ, i.e., for all primes ℓ there exists a

constant λ_ℓ such that

$$T_\ell(f) = \lambda_\ell f. \tag{16.5}$$

Then one gets relations

$$a_{\ell n} - \lambda_\ell a_n + \chi(\ell)\ell^{k-1}a_{n/\ell} = 0 \tag{16.6}$$

for all $n \geq 1$, and where we have set $a_{n/\ell} = 0$ if $\ell \not| n$. This relation gives $a_\ell = \lambda_\ell a_1$. Moreover if $a_1 = 0$ then one checks that $f = 0$. Thus we can assume a_1 is nonzero for any Hecke eigenform and we normalize any Hecke eigenform by setting $a_1 = 1$.

From the above discussion it is evident the Fourier coefficients of normalized Hecke eigenforms are the eigenvalues of the corresponding Hecke operators. This gives recursion relations between Fourier coefficients.

$$a_{\ell n} = a_\ell a_n - \chi(\ell)\ell^{k-1}a_{n/\ell} \tag{16.7}$$

and, in particular, if we take $n = \ell^m$ then we get

$$a_{\ell^{m+1}} = a_\ell a_{\ell^m} - \chi(\ell)\ell^{k-1}a_{\ell^{m-1}} \tag{16.8}$$

and $a_{rn} = a_r a_n$.

17. New forms

Let $M|N$ and let $d|(N/M)$ and assume $d > 1$ (so M divides N properly). Let $f \in S_k(\Gamma_1(M))$. Then the mapping $S_k(\Gamma_1(M)) \to S_k(\Gamma_1(N))$ defined by $f(z) \mapsto f(dz)$ is injective and this mapping takes forms with nebentype to forms with the same nebentype and it takes eigen vectors of Hecke operators T_ℓ for $\ell \not| N$ to eigenforms. We define $S_k(\Gamma_1(N))^{\text{old}}$ to be space of all forms which arise in this way from forms of lower level.

Definition 17.1. A form $f \in S_k(\Gamma_1(N))$ which is not in $S_k(\Gamma_1(N))^{\text{old}}$ is called a newform. The set of new forms is a subspace $S_k(\Gamma_1(N))$ and we denote it by $S_k(\Gamma_1(N))^{\text{new}}$.

It is a basic fact in the theory of new forms that $S_k(\Gamma_1(N))^{\text{new}}$ has a basis consisting of normalized new eigenforms. For a proof see [29].

Remark 17.2. Normalized Hecke newforms are uniquely identified by their Fourier coefficients and the Fourier coefficients are all algebraic numbers and that field K_f all the Fourier coefficients of a new eigen form f is a finite extension of $\mathbb{Q}$, i.e., K_f is a number field. For a proof see [42].

18. Galois representations associated to modular forms

Deligne's proof of the Ramanujan conjecture also produced the Galois representation associated to normalized new cusp eigenform. This representation was also constructed by Shimura for $k = 2$.

Theorem 18.1 (Deligne)**.** *Let $f \in S_k(\Gamma_0(N), \chi)$ be a normalized new eigenform and let $f = \sum_1^\infty a_n q^n$ be its Fourier expansion. Let $K_f = \mathbb{Q}(a_0, a_1, \cdots)$ be the field of Fourier coefficients of f and let $\mathfrak{p}$ be any prime lying over p in K_f. Let $K_{f,\mathfrak{p}}$ be the p-adic field associated to K_f at $\mathfrak{p}$. Then there exists a two dimensional, irreducible representation,*

$$\rho_{f,p} : G_{\mathbb{Q}} \to \mathrm{GL}_2(K_{f,\mathfrak{p}}) \tag{18.2}$$

with the following properties:

1. *$\rho_{f,p}$ is unramified outside pN,*
2. *for all $\ell \not| (pN)$ the characteristic polynomial of $\rho(\mathrm{Frob}_\ell)$ is given by the formula:*

$$\det(1 - X\rho(\mathrm{Frob}_\ell)) = 1 - a_\ell X + \ell^{k-1}\chi(\ell)X^2 \tag{18.3}$$

The following consequence of Theorem 12.11 and Proposition 13.18 shows that the representation constructed in Theorem 18.1 is characterized, up to isomorphism, by the properties listed in Theorem 18.1.

Proposition 18.4. *Any continuous irreducible, finite dimensional representation of $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ is determined, up to isomorphism, by the traces of Frobenius elements at all the primes where the representation is unramified.*

Remark 18.5. Such a representation $\rho_{f,p}$ has the property that $\det(\rho(\mathrm{Frob}_\ell)) = -1$

19. The Ramanujan Estimate

The following estimate for the size of the Fourier coefficients of normalized Hecke eigen cusp forms of level N, weight k and nebentype χ was conjectured by Ramanujan (see [35]) and Deligne (see [11]) showed that it is a consequence of the Weil conjectures (see [12]).

The estimate is a generalization of the Hasse-Weil estimate for elliptic curves (see Theorem 6.1).

Theorem 19.1 (Deligne; Ramanujan)**.** *Let $f = \sum_{n=1}^\infty a_n q^n$ be a normalized Hecke eigen form in $S_k(\Gamma_1(N))^{\mathrm{new}}$. Then for any prime ℓ we have*

$$|a_\ell| \leq 2\ell^{\frac{k-1}{2}}. \tag{19.2}$$

For $k = 2$, this estimate was also proved by Eichler and Shimura.

20. Reducing Galois representations modulo p

From now on we will assume that we have a normalized new Hecke eigen cusp form f such that $K_f = \mathbb{Q}$.

Most of theory outlined in this section works with out assumption. But we are restricting ourselves to this case as it keeps the notation fairly simple and transparent.

Let $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_n(\mathbb{Q}_p)$ be a continuous irreducible representation. As $G_{\mathbb{Q}}$ is a compact group, it is easy to see, using continuity of ρ, that the image of ρ is contained in a compact subgroup of the target. One can explicitly see this by observing that $G_{\mathbb{Q}}$ stabilizes a lattice $L \cong \mathbb{Z}_p^n \subset \mathbb{Q}_p^n$. Using this lattice we see that we ρ is the composite of $G_{\mathbb{Q}} \to \mathrm{GL}(L) \tilde{=} \mathrm{GL}_n(\mathbb{Z}_p) \hookrightarrow \mathrm{GL}_n(\mathbb{Q}_p)$. In other words we can arrange things in such a way that ρ has p-adic integer matrix entries. So we can then reduce these matrices modulo p. Thus we obtain, a representation $\bar{\rho} : G_{\mathbb{Q}} \to GL_n(\mathbb{Z}_p) \to \mathrm{GL}_n(\mathbb{Z}/p)$. Such a representation $\bar{\rho}$ is not in general unique and its construction depends in a rather strong way on the choice of the lattice which was used to carry out the reduction.

However, it is standard fact that such representation $\bar{\rho}$ is unique (up to isomorphism) if it is irreducible and its formation is independent of the choice of the lattice.

Example 20.1. Let $\Delta(q) = \sum_{n=1}^{\infty} \tau(n) q^n$ be the Ramanujan modular form of weight twelve and level 1. Take $p = 691$. Then Ramanujan observed that $\tau(n) \equiv \sigma_{11}(n) \mod 691$. This congruence implies that the mod 691 representation associated to Δ is reducible. For more on the connection between congruences between Fourier coefficients of modular forms and Galois representations see Serre's Seminar Bourbaki talk on the work of Swinnerton-Dyer [39].

21. Galois representations arising from elliptic curves

Let $E/\mathbb{Q}$ be an elliptic curve over $\mathbb{Q}$. Fix a prime p. One has a family of Galois representations associated to E, p, given by the action of the Galois group $G_{\mathbb{Q}}$ on the p^n-torsion points $E[p^n]$, which are denoted

$$\rho_{E,p^n} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{Z}/p^n) \tag{21.1}$$

It is not very difficult to verify that the composite homomorphism

$$\rho_{E,p^n} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{Z}/p^n) \to \mathrm{GL}_2(\mathbb{Z}/p^{n-1}) \tag{21.2}$$

is $\rho_{E,p^{n-1}}$. Such a system of representations can be assembled to give a continuous representation (note our abuse of notation)

$$\rho_{E,p} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{Z}_p). \tag{21.3}$$

Theorem 21.4. *Let $E/\mathbb{Q}$ be an elliptic curve and let N_E be its conductor. Then*

(1) *$\rho_{E,p}$ is unramified for primes not dividing pN_E and*
(2) *let $\ell \not| pN_E$ be a prime. Then*

$$\det(1 - X\rho_{E,p}(\mathrm{Frob}_\ell)) = 1 - a_\ell X + \ell X^2. \tag{21.5}$$

where $a_\ell = \ell + 1 - \#E(\mathbb{F}_\ell)$.

Remark 21.6. It is clear from the above theorem and Deligne's theorem 18.1 that the representation associated to an elliptic curves looks like the representation associated to a modular form of level N_E, weight 2 and $\chi = 1$.

22. The Shimura-Taniyama-Weil Conjecture

There are many equivalent definition of modularity of an elliptic curve. We will restrict our attention to the line of thought which emerged in the discussion at the end of previous section.

Definition 22.1. Let $E/\mathbb{Q}$ be an elliptic curve. We will say that E is a modular elliptic curve if the representation

$$\rho_{E,p} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{Z}_p) \tag{22.2}$$

if there exists a normalized new cusp form f on $\Gamma_1(N)$ such that

$$a_\ell(f) = a_\ell(E) \tag{22.3}$$

for all but finite number of primes (not dividing the level), and where $a_\ell(f)$ is the ℓ^{th} Fourier coefficient of f.

Remark 22.4. By Proposition 18.4, we observe that $\rho_{E,p}$ is isomorphic to the Galois representation of f at p. Thus to say that E is modular is equivalent to saying that $a_\ell(E) = \ell + 1 - \#(E(\mathbb{F}_\ell))$ are the eigenvalues of the Hecke operators on a normalized Hecke newform of some level, of weight two and $\chi = 1$.

It is a standard fact that if $\rho_{E,p}$ arises from a modular form f for one prime p then it does so for all primes p. Thus in this sense the definition is independent of the prime p.

Conjecture 22.5 (Shimura-Taniyama-Weil)**.** Every elliptic curve $E/\mathbb{Q}$ is modular.

This conjecture is now a theorem of Breuil, Conrad, Diamond and Taylor, Wiles [3] and its proof uses methods of [49], [46].

23. Serre's Conjecture

Serre in [41] formulated a mod p version of a similar conjecture. This conjecture of Serre still remains intractable.

In this section we will consider representations $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\bar{\mathbb{F}}_p)$. Assume that ρ is irreducible and continuous. Observe that continuity of ρ implies that the image is finite and hence is contained in $\mathrm{GL}_2(\mathbb{F}_q)$ for some finite extension $\mathbb{F}_q/\mathbb{F}_p$.

Definition 23.1. Let ρ be as above. We will say ρ is odd if $\det(\rho(c)) = -1 \in \bar{\mathbb{F}}_p^*$ where $c : \bar{\mathbb{Q}} \to \bar{\mathbb{Q}}$ is complex conjugation.

We remark that if $p = 2$ then $-1 = 1$ so there is no restriction at $p = 2$.

Definition 23.2. We will say that $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\bar{\mathbb{F}}_p)$ is modular if there exists a new eigen cusp form f on $\Gamma_1(N)$ for some $N \geq 1$ such that $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\bar{\mathbb{F}}_p^*)$ can be obtained by reduction the representation $\rho_{f,\mathfrak{p}} : G_{\mathbb{Q}} \to \mathrm{GL}_2(K_{f,\mathfrak{p}})$ modulo $\mathfrak{p}$ for some prime $\mathfrak{p}|p$ in K_f.

Conjecture 23.3 (Serre)**.** Let $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\bar{\mathbb{F}}_p)$ be any continuous, irreducible, odd representation of $G_{\mathbb{Q}}$. Then ρ is modular.

Remark 23.4. The condition that ρ is odd is necessary by Remark 18.5.

Serre also gave a recipe for computing the level, $N(\rho)$ and the weight $k(\rho)$ and the nebentype ϵ_ρ of such a form. As a consequence of [5], [13], [19], [32], [37] that if ρ is modular of some level weight and nebentype, then it does indeed arise from modular form of weight and level predicted by Serre.

Remark 23.5. For our purposes it is sufficient to know that for a representation $\rho : G_{\bar{\mathbb{Q}}} \to \mathrm{GL}_2(\bar{\mathbb{F}}_p)$ the predicted level $N(\rho)$ is coprime to p and is divisible only by prime $\ell \neq p$ which for which ρ is ramified.

24. Frey Elliptic Curves

Let A, B, C be pair wise coprime integers such that $A + B + C = 0$. Then we can associate an elliptic curve to such a triple of integers. Let

$$E_{A,B,C} : y^2 = x(x - A)(x + B) \tag{24.1}$$

For this curve we have $\Delta_E = 16(ABC)^2$. And hence as one of A, B, C is even we have

Lemma 24.2. *For elliptic curve $E_{A,B,C}$ we have $p|\Delta \Leftrightarrow p|ABC$.*

We now study the bad reduction of $E_{A,B,C}$ using Proposition 8.4. We observe that $c_4 = 16(A^2 + AB + B^2)$

Lemma 24.3. *Let $E = E_{A,B,C}$ be an elliptic curve associated to a triple of integers A, B, C as above. Then $E_{A,B,C}$ has good reduction outside primes p not dividing ABC. For any $p|ABC$, the curve E has semistable reduction modulo p.*

Proof. It is not difficult to verify that this equation is in fact minimal for all primes $p \neq 2$. The proof is easy for $p \neq 2$. for instance if $p|A$ then the reduction looks like $y^2 = x^2(x+B) \mod p$. (Note that $c_4 \not\equiv 0 \mod p$ as any $p|\Delta_E$ divides exactly one of A, B, C). For $p = 2$ we have to do a little more work. One essentially reduces to the case when $A \equiv -1 \mod 4$ and $B \equiv 0 \mod 32$. Then we can substitute

$$x = 4X \tag{24.4}$$

$$y = 8Y + 4X. \tag{24.5}$$

Then the equation of E reduces to

$$Y^2 + XY = X^3 + cX^2 + dX \tag{24.6}$$

where

$$c = \frac{B - 1 - A}{4} \tag{24.7}$$

$$d = -\frac{AB}{16} \tag{24.8}$$

and then the reduction of E modulo 2 is given by

$$Y^2 + XY = \begin{cases} X^3 & \text{if } A \equiv 7 \mod 8, \\ X^3 + X^2 & \text{if } A \equiv 3 \mod 8. \end{cases} \tag{24.9}$$

This curve has distinct tangents over $\bar{\mathbb{F}}_2$ and hence the reduction of E modulo 2 is semistable at 2. □

Thus we get

$$N_E = \prod_{p|ABC} p \tag{24.10}$$

$$j_E = \frac{2^8(C^2 - AB)^3}{A^2B^2C^2} \tag{24.11}$$

Proposition 24.12. *For $p \geq 5$ the representation*

$$\rho_{E,p} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{F}_p) \tag{24.13}$$

is irreducible and is unramified for primes not dividing ABC.

25. Frey Curves arising from FLT

Fix a prime $p \geq 5$. Let $a^p + b^p + c^p = 0$ be a nontrivial solution to Fermat's last theorem. We will assume that a, b, c are pairwise coprime and $abc \neq 0$. We will assume, without loss of generality, that $a \equiv -1 \mod 4$ and $b \equiv 0 \mod 4$. If we set $A = a^p, B = b^p, C = c^p$ then we get elliptic curves which were introduced and studied by G. Frey (see [20], [21], [22])

$$y^2 = x(x - a^p)(x + b^p) \tag{25.1}$$

26. Analysis of Ramification

In this section we analyze the ramification properties of the representations ρ obtained from the Frey elliptic curves $E = E_{A,B,C}$ for $A + B + C = 0$. Let $\rho = \rho_{E,p} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{F}_p)$ be the two dimensional representation of $G_{\mathbb{Q}}$ corresponding to its action on $E[p]$.

It suffices to concentrate on primes $\ell | (ABC)$ as ρ is unramified at primes p not dividing ABC. Fix a prime $\ell | ABC$. Then we know that E has semistable reduction at ℓ.

We will use Tate elliptic curves (see Section 11) to analyze the ramification at ℓ. For $\ell = 2$ we will need a refined version of Theorem 11.17 to take care of the possibility that the reduction type at 2 may be non-split semistable. But the difficulties are mostly technical and I will suppress this issue completely. The diligent reader is referred to [43].

By Tate's theorem, we can replace E by a Tate curve E_{q_ℓ} over $\mathbb{Q}_\ell$. Note that q_ℓ depends on E. In Remark 11.20 we had described $E_{q_\ell}[p]$ explicitly:

$$E_{q_\ell}[p] \cong \left\{ \zeta q_\ell^{m/p} \middle| \zeta^p = 1 \text{ and } m \in \mathbb{Z}/p \right\} \tag{26.1}$$

We can read of the ramification properties of ρ at ℓ from this description.

Proposition 26.2. *Let $\ell | ABC$. Then the extension corresponding to $E_{q_\ell}[p]$ is given by $\mathbb{Q}_\ell(\zeta, q_\ell^{1/p})$. In particular this extension is unramified for $\ell \neq 2, p$ if and only if $\nu_\ell(q_\ell) \equiv 0 \mod p$.*

Proof. If $\ell \neq 2, p$ and $\nu_\ell(q_\ell) \equiv 0 \mod p$ then $q_\ell = \ell^p u$ where u is unit in $\mathbb{Z}_\ell$. Thus the extension is given by $\mathbb{Q}_\ell(\zeta, u^{1/p})$ where ζ is a p^{th}-root of unity and $\ell \neq p$ and u is a unit in $\mathbb{Z}_\ell$. Thus this extension is unramified at ℓ. Conversely, if $\mathbb{Q}_\ell(\zeta, q_\ell^{1/p})$ is unramified at ℓ then we see that q_ℓ is a p^{th} power up to a unit in $\mathbb{Z}_\ell$. □

One can also carry out the analysis at $\ell = p$ however, the extension is ramified at $\ell = p$. But the ramification is fairly controlled (Serre calls this case peu ramifie) if and only if $\nu_p(q) \equiv 0 \mod p$. So we will make the following ad hoc definition.

Definition 26.3. Let $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{F}_p)$ be a representation. We will say that ρ is peu ramifie at p if the extension at p is given by $\mathbb{Q}_p(\zeta, u^{1/p})$ where ζ is a p^{th}-root of unity and $u \in \mathbb{Z}_p$ is a p-adic unit.

We now apply this analysis to the Frey curve and deduce:

Proposition 26.4. *Let $a^p+b^p+c^p = 0$ be a solution to Fermat's last theorem such that a, b, c are pairwise coprime and $abc \neq 0$ and let E be the Frey elliptic curve associated to it. Then the representation*

$$\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{F}_p) \tag{26.5}$$

is unramified for all prime $\ell \neq 2, p$ and for $\ell = p$ the representation is peu ramifie.

Proof. The is immediate from the fact that $\Delta = (abc)^p$ up to a power of 2 and $\nu_\ell(q) = \nu_\ell(\Delta_E)$ and hence the result. □

27. Fermat's Last Theorem

Theorem 27.1 (Serre)**.** *Serre's conjecture implies Fermat's Last Theorem.*

Proof. Serre's conjecture 23.3 implies that the representation $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{F}_p)$ associated to the corresponding Frey elliptic curve is modular of some level $N(\rho)$, weight $k(\rho)$ and nebentype χ. If we follow Serre's recipe (see Remark 23.5) for calculating these invariants we see that $N(\rho) = 2$ because $N(\rho)$ is divisible by only those primes $\ell \neq p$ for which ρ is ramified. Thus by Proposition 26.2 we see that $N(\rho) = 2$. The fact that ρ is peu ramifie at p gives that $k(\rho) = 2$ and recipe for nebentype gives and $\chi = 1$.

Thus the representation associated to the Frey curve arises from a newform of weight 2, level 2 and nebentype 1. But it is known (and fairly elementary to prove) that there are no such forms (see for instance [36]). Thus we arrive at contradiction. This proves the theorem. □

References

[1] B. Birch and H. P. F. Swinnerton Dyer. Notes on elliptic curves I. *J. Reine Angew. Math.*, 212:7–25, 1963.

[2] B. Birch and H. P. F. Swinnerton Dyer. Notes on elliptic curves II. *J. reine angew. Math.*, 218:79–108, 1965.

[3] C. Breuil, B. Conrad, F. Diamond, and R. Taylor. Modular elliptic curves over **Q**: Wild 3-adic exercises. *Preprint*, 2000.

[4] H. Carayol. Sur les représentations ℓ-adiques associées aux formes modulaires de Hilbert. *Ann. Sci. École Norm. Sup. (4)*, 19(3):409–468, 1986.

[5] H. Carayol. Sur les représentations galoisiennes modulo ℓ attachées aux formes modulaires. *Duke Math. J.*, 59(3):785–801, 1989.

[6] N. Chebotarev. Die bestimmung der dichtigkeit einer menge von primzahlen, welche zu einer gegebenen substitutionsklasse gehören. *Math. Ann.*, 95:151–228, 1925.

[7] B. Conrad. Ramified deformation problems. *Duke Math. J.*, 97(3):439–513, 1999.
[8] B. Conrad, F. Diamond, and R. Taylor. Modularity of certain potentially Barsotti-Tate Galois representations. *J. Amer. Math. Soc.*, 12(2):521–567, 1999.
[9] G. Cornell, J. Silverman, and G. Stevens, editors. *Modular forms and Fermat's last theorem*. Springer-Verlag, 1995.
[10] C. Curtis and I. Reiner. *Representation theory of finite groups and associated algebras*. Interscience, New-York-London, 1962.
[11] P. Deligne. Formes modulaires et représentations ℓ-adiques. In *Sem. Bourbaki*, volume 179 of *Lecture Notes in Mathematics*, Berlin, February 1969. Springer-Verlag.
[12] P. Deligne. La conjecture de Weil I. *Publ. Math. I.H.E.S*, 43:5–77, 1974.
[13] F. Diamond. The refined conjecture of Serre. In *Elliptic curves, modular forms and Fermat's Last Theorem*, volume 1 of *Number Theory*, pages 22–37, Cambridge, MA., 1993. International Press.
[14] F. Diamond. On deformation rings and Hecke rings. *Ann. of Math. (2)*, 144(1):137–166, 1996.
[15] F. Diamond. The Taylor-Wiles construction and multiplicity one. *Invent. Math.*, 128(2):379–391, 1997.
[16] F. Diamond, H. Darmon, and R. Taylor. *Fermat's last theorem*, pages 1–154. Current developments in mathematics. International Press, Cambridge, MA, 1997.
[17] F. Diamond and K. Kramer. Modularity of a family of elliptic curves. *Math. Res. Let.*, 2(3):299–304, 1995.
[18] B. Dwork. On the rationality of the zeta function of an algebraic variety. *Amer. Jour. of Math.*, 82:631–648, 1960.
[19] B. Edixhoven. The weight in serre's conjectures on modular forms. *Invent. Math.*, 109(3):563–594, 1992.
[20] G. Frey. Rationale Punkte auf Fermatkurven und getwisteten Modulkurven. *Journal reine angew. Math.*, 331:185–191, 1982.
[21] G. Frey. Links between stable elliptic curves and certain Diophantine equations. *Ann. Univ. Saraviensis, Ser. Math.*, 1:1–40, 1986.
[22] G. Frey. Links between elliptic curves and solutions of $A - B = C$. *J. Indian Math. Soc. N.S.*, 51:117–145, 1987.
[23] H. Hasse. *Mathematische Abhandlungen*, volume 1-3. Walter de Gruyter, Berlin-New-York, 1975.
[24] H. Hida. Iwasawa modules attached to congruences of cusp forms. *Ann. Scient. de l'E. N.S.*, 19:231–273, 1986.
[25] H. Hida. *Elementary theory of L-functions and Eisenstein series*, volume 26 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 1993.
[26] K. Ireland and M. Rosen. *A classical introduction to modern number theory*, volume 84 of *Graduate Texts in Mathematics*. Springer-Verlag, New-York, second Edition edition, 1990.
[27] N. Katz. An overview of Deligne's proof. In *Proc. Symp. Pure Math.*, volume 28. AMS, 1976.
[28] N. Koblitz. *A course in number theory and cryptography*, volume 114 of *Graduate Texts in Mathematics*. Springer-Verlag, New-York, second Edition edition, 1994.
[29] S. Lang. *Introduction to Modular forms*, volume 222 of *Grundlehren de Mathematishen Wissenschaften*. Springer-Verlag, Berlin, corrected and reprint 1995 edition, 1976.
[30] M. Laska. An algorithm for finding a minimal Weierstrass equation for an elliptic curve. *Math. Comp.*, 38:257–260, 1982.

[31] S. Lubkin. A p-adic proof of Weil's conjectures. *Ann. of Math. (2)*, 87:195–255, 1968.
[32] B. Mazur. Modular curves and the Eisenstein ideal. *Publ. Math. IHES*, 47:33–186, 1977.
[33] B. Mazur and A. Wiles. On p-adic analytic families of modular forms. *Comp. Math.*, 59:231–264, 1986.
[34] Barry Mazur. Deforming Galois Representations. In Y. Ihara, K. Ribet, and J.-P. Serre, editors, *Galois groups over* $\mathbf{Q}$, number 16 in Mathematical Sciences Research Institute Publications, pages 385–438, Berlin, 1989. Springer-Verlag.
[35] S. Ramanujan. *Collected Mathematical Papers.*
[36] R. Rankin. *Modular forms and functions.* Cambridge University Press, Cambridge, 1977.
[37] K. Ribet. On modular representations of $Gal(\bar{\mathbf{Q}}/\mathbf{Q})$ arising from mdoular forms. *Invent. Math.*, 100(2):431–476, 1990.
[38] J.-P. Serre. *Abelian ℓ-adic representations.* Benjamin, New York.
[39] J.-P. Serre. Congruences et formes modulaires (d'aprés h. p. f. Swinnerton-Dyer). In *Sém. Bourbaki*, number 416, 1971/72.
[40] J.-P. Serre. *Local fields*, volume 67 of *Graduate Texts in Mathematics.* Springer, Berlin, Springer-Verlag, 1979.
[41] J.-P. Serre. Sur les représentations modulaires de degré 2 de $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$. *Duke Math. Journal*, 54(1):179–230, 1987.
[42] G. Shimura. *Introduction to arithmetic theory of automorphic forms.* Princeton University Press, 1971.
[43] Joseph Silverman. *The arithmetic of elliptic curves*, volume 106 of *Graduate Text in Mathematics.* Springer-Verlag, Berlin, 1985.
[44] J. Tate. The arithmetic of elliptic curves. *Invent. Math.*, 23:179–206, 1974.
[45] J. Tate. Algorithm for determining the type of a singular fibre in an elliptic pencil. In *Modular functions of one variable IV*, volume 476 of *Lecture Notes in Mathematics*, pages 33–52, Berlin, 1975. Springer-Verlag.
[46] R. Taylor and A. Wiles. Ring theoretic properties of certain Hecke algebras. *Ann. of Math. (2)*, 141:553–572, 1995.
[47] André Weil. Sur un théorème de Mordell. *Bull. Sci. Math.*, 54:182–191, 1930.
[48] André Weil. Number of solutions of equations in finite fields. *Bull. AMS*, 55:497–508, 1949.
[49] A. Wiles. Modular elliptic curves and Fermat's last theorem. *Ann. of Math. (2)*, 141:443–551, 1995.

Kirti Joshi, Department of Mathematics, University of Arizona, Tucson, AZ 85721

E-mail address: kirti@math.arizona.edu