CHAPTER 12

Elliptic curves (II)

In this chapter we discuss the arithmetic of elliptic curves. In §12.1 we show the significance of the conjecture known as the Taniyama-Shimura-Weil conjecture in the arithmetic of elliptic curves. In §12.2 we give a very rough idea of the proof of Fermat's Last Theorem by Wiles (Wiles and Taylor-Wiles). Because of the level and the aim of this book, we state most theorems in this chapter without proof. An intrigued reader should consult other books.

12.1. Elliptic curves over the rational number field

(a) Rational points over finite fields. We first consider polynomial equations of degree 3 with integer coefficients. For example, take

$$y^2 = x^3 - x.$$

We have already discussed the integral and rational solutions of this equation. Here, we consider an easier problem: For a prime number l, count the number of solutions in $\mathbb{F}_l = \mathbb{Z}/l\mathbb{Z}$ of the equation modulo l. We regard this problem as easy because for a given prime number, finding solutions can be done, in principle, by a finite number of steps of computations. Let us actually find the solutions. We have the following list.

$\mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z}$	$(0,0),(1,0)$
$\mathbb{F}_3 = \mathbb{Z}/3\mathbb{Z}$	$(0,0),(1,0),(2,0)$
$\mathbb{F}_5 = \mathbb{Z}/5\mathbb{Z}$	$(0,0),(1,0),(2,1),(2,4),(3,2),(3,3),(4,0)$
$\mathbb{F}_7 = \mathbb{Z}/7\mathbb{Z}$	$(0,0),(1,0),(4,2),(4,5),(5,1),(5,6),(6,0)$

Table 12.1 shows the number of solutions only.

The reader should pretend as if s/he were Fermat or Gauss and try to find the pattern of the number of solutions.

prime number l	2	3	5	7	11	13	17	19	23	29	31	37	41	43	47	53	59	61	67
number of solutions	2	3	7	7	11	7	15	19	23	39	31	39	31	43	47	39	59	71	67

Table 12.1

Question 1. Show that the number of solutions is congruent to 3 modulo 4, except for the case of $l = 2$.

As a matter of fact, we have the following.

Theorem 12.1 (Gauss). *Let l be an odd prime number.*

(1) *If $l \equiv 3 \bmod 4$, then the number of solutions in $\mathbb{F}_l$ to the equation $y^2 = x^3 - x$ equals l.*

(2) *If $l \equiv 1 \bmod 4$, then, as we stated in Proposition 0.2 in Number Theory 1, l can be written in the form*

$$l = a^2 + b^2, \quad a, b \in \mathbb{Z}.$$

We choose a as an odd number, b an even number, and we choose the sign of a so that we have

$$\begin{aligned} a &\equiv 1 \bmod 4 \quad \text{if } \ b \equiv 0 \bmod 4, \\ a &\equiv 3 \bmod 4 \quad \text{if } \ b \equiv 2 \bmod 4. \end{aligned}$$

Then, the number of solutions in $\mathbb{F}_l$ to the equation $y^2 = x^3 - x$ equals $l - 2a$.

What we want to note here is that there exists such a law for the number of solutions. The number of solutions could be at random for each prime number, but we have such a beautiful theorem. Is there any reason for this? Do such properties exist for a more general setting?

Let us consider a slightly different equation

$$y^2 + y = x^3 - x^2.$$

Similar calculations show the following.

$\mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z}$	$(0,0), (0,1), (1,0), (1,1)$
$\mathbb{F}_3 = \mathbb{Z}/3\mathbb{Z}$	$(0,0), (0,2), (1,0), (1,2)$
$\mathbb{F}_5 = \mathbb{Z}/5\mathbb{Z}$	$(0,0), (0,4), (1,0), (1,4)$
$\mathbb{F}_7 = \mathbb{Z}/7\mathbb{Z}$	$(0,0), (0,6), (1,0), (1,6), (4,2), (4,4), (5,1), (5,5), (6,3)$

THEOREM 12.2 (Eichler). *For a positive integer n, let a_n be the integer coefficient of q^n in the power series*

$$q\prod_{n=1}^{\infty}(1-q^n)^2(1-q^{11n})^2$$
$$= q - 2q^2 - q^3 + 2q^4 + q^5 + 2q^6 - 2q^7 + \cdots = \sum_{n=1}^{\infty} a_n q^n.$$

Then, for $l \neq 11$, the number of solutions to the equation $y^2 + y = x^3 - x^2$ in $\mathbb{F}_l$ equals $l - a_l$.

Note that the above power series is a modular form of weight 2 and level 11 (§9.1(d)). If we write

$$q\prod_{n=1}^{\infty}(1-q^{4n})^2(1-q^{8n})^2 = \sum_{n=1}^{\infty} b_n q^n,$$

we can prove that the number of solutions to the equation $y^2 = x^3 - x$ in $\mathbb{F}_l$ equals $l - b_l$ (see Theorem 12.1).

(b) Reduction mod *l*. Equations such as those above can be considered as the equations of elliptic curves. Here, for a given elliptic curve, we systematically study the curves over $\mathbb{F}_l$ obtained by reducing its coefficients mod l.

If the reduction mod l of the equation of an elliptic curve with integer coefficients is an elliptic curve over $\mathbb{F}_l$, namely, it can be converted over $\mathbb{F}_l$ to the form $y^2 = f(x)$ where $f(x)$ does not have a multiple root, then we say that it has *good reduction* at l. Otherwise it has *bad reduction* at l. Consider $y^2 = x^3 - x$, for example. Since $x^3 - x = x(x-1)(x+1)$, and 0, ± 1 are all distinct in $\mathbb{F}_l$ when $l \neq 2$, it has a good reduction at any odd prime number. However, the above definition is not quite precise, and we must state it more accurately.

In Chapter 1 of *Number Theory 1*, an elliptic curve over $\mathbb{Q}$ is defined as the curve given by an equation of the form

$$y^2 = ax^3 + bx^2 + cx + d \quad (a, b, c, d \in \mathbb{Q})$$
$$a \neq 0, \quad \text{the right-hand side has no multiple root}$$

(see §1.1(b) of *Number Theory 1*). For example, if we replace $3x$ by x in $y^2 = 27x^3 - 3x$, we obtain $y^2 = x^3 - x$. In other words they are isomorphic over $\mathbb{Q}$. However, when the coefficients are reduced mod 3, the former equation becomes $y^2 = 0$, which is no longer an elliptic curve over $\mathbb{F}_3$. It is troublesome that the behavior of the reduction mod l differs when we choose a different equation of the same curve. We deal with this problem as follows.

We first define an elliptic curve E over the integers as the curve defined by the cubic equation of the form

$$(12.1) \quad E : y^2 + a_1xy + a_3 = x^3 + a_2x^2 + a_4x + a_6, \quad a_1, \dots, a_6 \in \mathbb{Z}$$

that satisfies the condition (12.2) below. If we define

$$\begin{aligned} b_2 &= a_1^2 + 4a_2, \\ b_4 &= 2a_4 + a_1a_3, \\ b_6 &= a_3^2 + 4a_6, \end{aligned}$$

and replace y by $\frac{1}{2}(y - a_1x - a_3)$, then the equation (12.1) can be transformed to

$$y^2 = 4x^3 + b_2x^2 + 2b_4x + b_6.$$

Now define $b_8 = a_1^2a_6 + 4a_2a_6 - a_1a_3a_4 + a_2a_3^2 - a_4^2$, and define

$$\Delta = -b_2^2b_8 - 8b_4^3 - 27b_6^2 + 9b_2b_4b_6.$$

We call Δ the discriminant of E. The curve E given by (12.1) is called an elliptic curve over the integers if the condition

$$(12.2) \qquad \Delta \neq 0$$

is satisfied.

If we take u, r, s, $t \in \mathbb{Q}$ $(u \neq 0)$, and replace x by $u^2x + r$, and y by $u^3y + su^2x + t$ in (12.1), we obtain another equation of the same form. (It can be proved that any two elliptic curves isomorphic over $\mathbb{Q}$ can transform each other by this change of variables.) Suppose the new equation obtained in this way also has integer coefficients. Among all such equations of an elliptic curve with integer coefficients, the one whose absolute value of the discriminant is minimal, is called a *minimal Weierstrass model.*

For example, $y^2 = x^3 - x$ and $y^2 + y = x^3 - x^2$ are both minimal Weierstrass models. (The discriminants are 32 and -11, respectively.)

Now, let E be an elliptic curve defined over $\mathbb{Q}$. Using the transformation described above, we transform it to a minimal Weierstrass model, and we compute b_i and Δ. If a prime number l does not divide Δ, then we say that E has *good reduction* at l. Otherwise we say that E has *bad reduction* at l. If E has good reduction at l, then the equation obtained by reducing (12.1) mod l becomes an elliptic curve over $\mathbb{F}_l$. The number of prime numbers at which E has bad reduction is finite since such prime numbers divide Δ.

Suppose E has bad reduction at l. If l does not divide $b_2^2 - 24b_4$, then we say that E has *multiplicative reduction* at l. If l divides $b_2^2 - 24b_4$, then we say that E has *additive reduction* at l. If E has good or multiplicative reduction at l, we say that E has *semi-stable reduction* at l. If E has semi-stable reduction at all prime numbers, that is, it has either good or multiplicative reduction at all prime numbers, then we say that E is a *semi-stable elliptic curve*.

If E has multiplicative reduction at l, then E mod l has a singular point (a point at which both $\frac{\partial f}{\partial x}$ mod l and $\frac{\partial f}{\partial y}$ mod l becomes 0, where $f = y^2 + a_1xy + a_3y - x^3 - a_2x^2 - a_4x - a_6$). Consider the equation of E as an equation over $\mathbb{F}_l$, and consider tangent lines formally. If the slopes of the tangent lines at the double point belong to $\mathbb{F}_l$, then we say that E has *split multiplicative reduction* at l. Otherwise E has *nonsplit multiplicative reduction* at l.

For example, $y^2 = x^3 - x$ has additive reduction at $l = 2$, and it has good reduction at all other primes. $y^2 + y = x^3 - x^2$ has good reduction except at 11, and it has split multiplicative reduction at $l = 11$.

(c) *n*-torsion points and an action of Galois group. Let E be an elliptic curve defined over $\mathbb{Q}$. The graph formed by the set of all real solutions of an elliptic curve was described in Chapter 1 of *Number Theory 1*. The graph formed by the set of all complex solutions (a Riemann surface, to be precise) is a torus (the surface of a doughnut with one hole). This follows from the following fact. The equation (12.1) can be transformed to the form

$$y^2 = 4x^3 - g_2x - g_3$$

over the complex numbers, and we have the following fact.

THEOREM 12.3. *For an elliptic curve $E : y^2 = 4x^3 - g_2x - g_3$, there exist $\mathbb{R}$-linearly independent complex numbers ω_1 and ω_2 satisfying the following conditions*

(i) $$g_2 = 60 \sum_{(m,n)\neq(0,0)} \frac{1}{(m\omega_1 + n\omega_2)^4},$$
$$g_3 = 140 \sum_{(m,n)\neq(0,0)} \frac{1}{(m\omega_1 + n\omega_2)^6},$$
where (m, n) runs over all pairs of integers except for $(0, 0)$.

(ii) *If we define*
$$\wp(z) = \frac{1}{z^2} + \sum_{(m,n)\neq(0,0)} \left(\frac{1}{(z - m\omega_1 - n\omega_2)^2} - \frac{1}{(m\omega_1 + n\omega_2)^2} \right),$$
and $L = \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2$, then the map
$$\begin{aligned} \mathbb{C}/L &\xrightarrow{\simeq} E(\mathbb{C}) \\ z &\mapsto (\wp(z), \wp'(z)) \end{aligned}$$
is an isomorphism of groups. Here, $E(\mathbb{C})$ is the abelian group formed by all the $\mathbb{C}$-rational points of E (see §1.2(a)).

Topologically, $\mathbb{C}/L$ is homeomorphic to a torus. For a positive integer n, the subgroup $E[n]$ of all n-torsion points, that is, the elements that vanish by the multiplication by n, is isomorphic to $\mathbb{Z}/n\mathbb{Z} \oplus \mathbb{Z}/n\mathbb{Z}$:
$$E[n] = \{x \in \mathbb{C}/L \mid nx = 0\} \simeq \mathbb{Z}/n\mathbb{Z} \oplus \mathbb{Z}/n\mathbb{Z}.$$
Suppose again that E is defined over $\mathbb{Q}$. If $nP = 0$ for $P = (x, y) \in E(\mathbb{C})$, then it can be proved that both x and y coordinates of P are algebraic over $\mathbb{Q}$.

Let $\overline{\mathbb{Q}}$ be an algebraic closure of the rational number field $\mathbb{Q}$, and let $G_{\mathbb{Q}} = \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ be its Galois group. For a point $P = (x, y) \in E[n]$, we define the action of an element $\sigma \in G_{\mathbb{Q}}$ by $\sigma(P) = (\sigma(x), \sigma(y))$. Corresponding to the above isomorphism, there exists a basis e_1, e_2 such that any $P \in E[n]$ can be expressed as
$$P = ae_1 + be_2, \quad a, b \in \mathbb{Z}/n\mathbb{Z}.$$
If $nP = 0$, then we have $n\sigma(P) = \sigma(nP) = 0$, and $\sigma(P)$ also belongs to $E[n]$. Thus, if we write
$$\sigma(e_1) = ae_1 + ce_2, \quad \sigma(e_2) = be_1 + de_2,$$

then the action of σ on the n-torsion points can be expressed by a matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL_2(\mathbb{Z}/n\mathbb{Z}).$$

In this way we can define a homomorphism of groups

$$G_{\mathbb{Q}} \to GL_2(\mathbb{Z}/n\mathbb{Z})$$
$$\sigma \mapsto \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

All these are summarized as Theorem 12.4(1) below. The above homomorphism satisfies the property in (2).

Theorem 12.4. (1) *Let E be an elliptic curve defined over $\mathbb{Q}$, $E[n]$ the subgroup of n-torsion points. Then, as an abelian group, $E[n]$ is isomorphic to $\mathbb{Z}/n\mathbb{Z} \oplus \mathbb{Z}/n\mathbb{Z}$. The coordinates of points in $E[n]$ are algebraic over $\mathbb{Q}$, and thus $G_{\mathbb{Q}} = \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ acts on $E[n]$. This action defines a homomorphism*

$$\rho_{E[n]} : G_{\mathbb{Q}} \to GL_2(\mathbb{Z}/n\mathbb{Z}).$$

(2) *Let $K_n/\mathbb{Q}$ be the extension corresponding to the kernel of the above homomorphism. Namely, K_n is the field such that $\mathrm{Gal}(K_n/\mathbb{Q}) = \mathrm{Ker}\,\rho_{E[n]}$. Suppose E has good reduction at a prime number l that is prime to n. Then, l is unramified in $K_n/\mathbb{Q}$.*

(d) Tate module. Let E be as above, and p a prime number. For a positive integer n, consider the subgroup $E[p^n]$ of p^n-torsion points. The projective limit with respect to the multiplication-by-p map $p : E[p^{n+1}] \to E[p^n]$

$$T_p(E) = \varprojlim E[p^n]$$

is called the *Tate module*. $T_p(E)$ is a free $\mathbb{Z}_p$-module of rank 2. The Galois group $G_{\mathbb{Q}}$ acts on $T_p(E)$. Let e_1, e_2 be a basis of $T_p(E)$ as a $\mathbb{Z}_p$-module, and write

$$\sigma(e_1) = ae_1 + ce_2, \quad \sigma(e_2) = be_1 + de_2.$$

Then, we obtain a continuous homomorphism

$$\rho_p : G_{\mathbb{Q}} \to GL_2(\mathbb{Z}_p)$$

by defining

$$\rho(\sigma) = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL_2(\mathbb{Z}_p).$$

This homomorphism may be considered as the projective limit of the homomorphisms $\rho_{E[p^n]} : G_{\mathbb{Q}} \to GL_2(\mathbb{Z}/p^n\mathbb{Z})$ coming from the action on $E[p^n]$.

We have the following relation between the number of solutions over finite fields of an equation and the action of the Galois group.

THEOREM 12.5. *Let $K_{p^\infty}/\mathbb{Q}$ be the extension corresponding to the kernel* $\operatorname{Ker}\rho_p$ *of the homomorphism*

$$\rho_p : G_{\mathbb{Q}} \to GL_2(\mathbb{Z}_p).$$

In other words, let K_{p^∞} be the extension satisfying $\operatorname{Gal}(\overline{\mathbb{Q}}/K_{p^\infty}) = \operatorname{Ker}\rho_p$. *Let l be a prime number different from p, and suppose E has good reduction at l. Then, l is unramified in $K_{p^\infty}/\mathbb{Q}$, and the Frobenius conjugacy class* Frob_l *satisfies*

$$\det(\rho_p(\operatorname{Frob}_l)) = l.$$

Thus, $\det(\rho_p)$ *coincides with the cyclotomic character κ* (§10.1(e)). *Furthermore, if we put*

$$\operatorname{Tr}(\rho_p(\operatorname{Frob}_l)) = a_l,$$

then a_l is an integer satisfying

$$\#E(\mathbb{F}_l) = l + 1 - a_l,$$

where $E(\mathbb{F}_l)$ is the group of all $\mathbb{F}_l$-rational points of E (the group consisting of all solutions of the equation in $\mathbb{F}_l$ and the origin). Thus, $\#E(\mathbb{F}_l)$ is the number of solutions plus 1.

It is rather surprising that the number of solutions of an equation can be obtained through the action of the Galois group. The above theorem also claims that $\operatorname{Tr}(\rho_p(\operatorname{Frob}_l))$ $(= a_l)$ does not depend on p. Thus, E determines a sequence $(a_l)_l$, where l runs through all the primes at which E has good reduction. $(a_l)_l$ turns out to be quite significant to the elliptic curve E.

(e) ζ function and L-function of an elliptic curve. In §7.4 of *Number Theory 2*, we gave an account for the Hasse ζ-function of a finitely generated ring over $\mathbb{Z}$. Similarly we can define Hasse ζ function for a scheme of finite type over $\mathbb{Z}$.

For an elliptic curve E over $\mathbb{Q}$, using the minimal Weierstrass model, the Hasse ζ function can be defined by

$$\begin{aligned}\zeta_E(s) = &\prod_{l:\text{good}} \frac{1 - a_l l^{-s} + l^{l-2s}}{(1-l^{1-s})(1-l^{-s})} \prod_{l:\text{split mult.}} \frac{1}{1-l^{1-s}} \\ &\times \prod_{l:\text{nonsplit mult.}} \frac{1+l^{-s}}{(1-l^{1-s})(1-l^{-s})} \prod_{l:\text{add.}} \frac{1}{(1-l^{1-s})(1-l^{-s})},\end{aligned}$$

where the first product runs over all primes at which E has good reduction, the second, third, and fourth products run over primes of split multiplicative, nonsplit multiplicative, and additive reduction, respectively. a_l in the first product is given by $a_l = l + 1 - \#E(\mathbb{F}_l)$. Since there are only finitely many primes at which E has bad reduction, the first product is the principal term. Notice that the sequence (a_l) arises here once again.

Define *L-function* of E by

$$L(E,s) = \prod_{l:\text{good}} \frac{1}{1 - a_l l^{-s} + l^{l-2s}} \times \begin{bmatrix}\text{contribution from} \\ \text{bad primes}\end{bmatrix}$$

$$\begin{bmatrix}\text{contribution from} \\ \text{bad primes}\end{bmatrix} = \prod_{l:\text{split mult.}} \frac{1}{1+l^{-s}} \prod_{l:\text{nonsplit mult.}} \frac{1}{1-l^{-s}}.$$

Then, we have

$$\zeta_E(s) = \zeta_{\mathbb{Z}}(s)\,\zeta_{\mathbb{Z}}(s-1)\,L(E,s)^{-1},$$

where $\zeta_{\mathbb{Z}}(s)$ is the Riemann ζ function. We may consider that the principal term of $\zeta_E(s)$ is $L(E,s)$. It is known that $L(E,s)$ converges absolutely for $\mathrm{Re}(s) > 3/2$. Hasse conjectured that $L(E,s)$ admits an analytic continuation to the entire complex plane. This conjecture was proved for the semi-stable case as a consequence of Wiles' theorem, which we describe below, and later his method was generalized. The conjecture has been solved completely now.

The L-function $L(E,s)$ is a very important function for an elliptic curve E. For example, there is a famous unsolved problem called the *Birch-Swinnerton-Dyer conjecture*. We cannot describe its complete version. We just state it as the relation between the rank of the Mordell-Weil group $E(\mathbb{Q})$ as an abelian group and the L-function. Birch and Swinnerton-Dyer conjectured that

$$\mathrm{ord}_{s=1} L(E,s) = \mathrm{rank}\, E(\mathbb{Q}).$$

In other words, an arithmetically important number, the rank of the Mordell-Weil group, can be extracted from the information of the L-function. In the case where the left-hand side equals 0 or 1, significant progress has been made recently. However, it is still an open problem when the left-hand side is greater than 1.

Recall that for a number field, the order of $\zeta_K(s)$ at $s = 0$ satisfies

$$\operatorname{ord}_{s=0} \zeta_K(s) = r_1 + r_2 - 1 = \operatorname{rank} \mathcal{O}_K^\times,$$

where $\mathcal{O}_K^\times$ is the unit group of K (see Theorem 7.10 in §7.2(a) of *Number Theory 2*). The Birch-Swinnerton-Dyer conjecture is an analog of this fact.

Let us summarize what we have described so far. For an elliptic curve defined over $\mathbb{Q}$, let $\mathcal{L}$ be the set of prime numbers at which E has good reduction. Then, a sequence $(a_l)_{l\in\mathcal{L}}$ of integers is defined. The integer a_l appears in the following important situations:

(1) If $\mathbb{F}_l$ is the finite field of l elements, the number of $\mathbb{F}_l$-rational points of E satisfies

$$a_l = l + 1 - \#E(\mathbb{F}_l).$$

(This can be regarded as the definition of a_l.)

(2) The action of the Galois group $G_\mathbb{Q}$ on the Tate module $T_p(E)$ induces a continuous homomorphism

$$\rho_p : G_\mathbb{Q} \to GL_2(\mathbb{Z}_p),$$

and we have

$$a_l = \operatorname{Tr}(\rho_p(\operatorname{Frob}_l))$$

for $l \neq p$.

(3) $(a_l)_{l\in\mathcal{L}}$ appears in the Euler products of ζ function and L-function of E.

(f) Modular elliptic curves. So far, we have described that it is very important to know the sequence $(a_l)_{l\in\mathcal{L}}$ for the arithmetic of an elliptic curve. Surprisingly, the number a_l is related to the world of modular forms.

CONJECTURE 12.6 (Taniyama-Shimura-Weil Conjecture). *Let E be an elliptic curve defined over $\mathbb{Q}$, $\mathcal{L}$ the set of prime numbers at which E has good reduction, and $(a_l)_l$ as above. Then there exists a modular form $\sum a_n q^n$ whose q^l coefficient of the q-expansion coincides with a_l.*

More precisely, for an elliptic curve defined over $\mathbb{Q}$, a positive integer N, called a conductor, is defined, and there is a modular form $\sum a_n q^n$ of weight 2 and level N (or level $\Gamma_0(N)$) that is an eigenfunction of all Hecke operators.

Theorem 12.2 states that this conjecture holds for the elliptic curve $y^2+y = x^3-x^2$. Also, as we stated at the end of (a), $y^2 = x^3-x$ is related to a modular form.

Taniyama, in 1955, stated a preliminary version of this conjecture as a problem. In the 1960s Shimura formulated the conjecture rigorously into the above form. Due to Weil's work in 1967, this conjecture became widely known.

In general, for a cusp form $f = \sum a_n q^n$ of weight 2 and level N with $a_1 = 1$ that is an eigenfunction of all Hecke operators, define $K = \mathbb{Q}_p(\{a_n \,|\, n \geq 2\})$. Then, Eichler and Shimura showed that $K/\mathbb{Q}_p$ is a finite extension, and there exists a continuous representation

$$\rho_f : G_{\mathbb{Q}} = \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to GL_2(K)$$

such that $\mathrm{Tr}(\rho_f(\mathrm{Frob}_l)) = a_l$ and $\det(\rho_f(\mathrm{Frob}_l)) = l$. This representation is obtained by decomposing the Tate module of the Jacobian $J_0(N)$ of the modular curve $X_0(N)$ by the action of Hecke operators and considering the $G_{\mathbb{Q}}$ action of its f-factor.

Suppose now that $f = \sum a_n q^n$ is a cusp form of weight 2 and level N that is an eigenfunction of all Hecke operators such that all the coefficients a_n are rational numbers with $a_1 = 1$. Then, in the decomposition of the Jacobian $J_0(N)$ of the modular curve $X_0(N)$ by the Hecke operators, the f-factor becomes an abelian variety of dimension 1, that is, an elliptic curve. Let E be this elliptic curve, and we have

$$\mathrm{Tr}(\rho_f(\mathrm{Frob}_l)) = \mathrm{Tr}(\rho_E(\mathrm{Frob}_l)) = a_l.$$

Thus, such an E satisfies the condition of Conjecture 12.6.

An elliptic curve obtained in this way and elliptic curves isogenous to it are called *modular elliptic curves*. Using Faltings' isogeny theorem proved generally, the Taniyama-Shimura-Weil conjecture states that all elliptic curves defined over $\mathbb{Q}$ are modular elliptic curves. The world of modular forms is a beautiful world full of symmetries. It is rather surprising at first that all elliptic curves come from such a world. For example, for a modular elliptic curve, its L-function is known to have an analytic continuation to the entire plane, which is

Hasse's conjecture, mentioned in (e). (See also Theorem 9.7.) Conversely, Weil showed, in 1967, that if the L-function $L(E, s)$ of an elliptic curve E defined over $\mathbb{Q}$ has an analytic continuation to the entire complex plane and it satisfies certain functional equations, E is modular. After Weil's work, Conjecture 12.6 became widely accepted as true.

Let $L(E, s)$ be the L-function of an elliptic curve defined over $\mathbb{Q}$. As we stated in (e), $L(E, s)$ is defined as the Euler product. Write it formally as the form of Dirichlet series

$$L(E, s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s}.$$

If n equals a prime number l, then the a_n on the right-hand side coincides with a_l in the defining Euler product of L-function (and thus, the a_l we have been discussing repeatedly). With this notation the Taniyama-Shimura-Weil conjecture can be stated as follows.

Conjecture 12.6$'$ (Taniyama-Shimura-Weil). *Let E be an elliptic curve defined over $\mathbb{Q}$ with conductor N, and*

$$L(E, s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s}$$

its L-function. Then

$$f = \sum_{n=1}^{\infty} a_n q^n$$

is a modular form of weight 2 *and level N.*

In the above statement we used ζ function to rewrite the conjecture. Let us give another version in terms of algebraic geometry. As we have already mentioned above, the conjecture is equivalent to the statement that any elliptic curve defined over $\mathbb{Q}$ appears as a factor of the Jacobian of a modular curve (more precisely, isogenous to a factor of the Jacobian). We can prove this is equivalent to the following.

Conjecture 12.6$''$ (Taniyama-Shimura-Weil). *Let E be an elliptic curve defined over $\mathbb{Q}$ with conductor N. Then there exists a nonzero morphism from a modular curve*

$$X_0(N) \to E.$$

In other words, E is parametrized by $X_0(N)$. It is interesting to compare this conjecture with a theorem of Gauss (see Proposition 5.13

in *Number Theory 2*), which states that any quadratic extension is contained in a cyclotomic field. An elliptic curve is a double cover of a projective line. The Taniyama-Shimura-Weil conjecture states that such a curve is always parametrized by a modular curve. In this way the Taniyama-Shimura-Weil conjecture is considered as an algebraic geometry analogue of Gauss' theorem.

We will state a version of the Taniyama-Shimura-Weil conjecture from the point of view of Galois representation in §12.2(c).

12.2. Fermat's Last Theorem

(a) Frey curve. Now we will discuss Fermat's Last Theorem. As a preparation, consider the following elliptic curve. Let A, B, C be relatively prime integers satisfying

$$\begin{aligned}&A+B=C,\\&A\equiv 3 \ (\mathrm{mod} 4), \quad B\equiv 0 \ (\mathrm{mod} 32).\end{aligned}$$

Consider the elliptic curve E defined by

$$E : y^2 = x(x-B)(x-C).$$

By a change of variables $x = 4X$, $y = 8Y + 4X$, the equation of E becomes

$$Y^2 + XY = X^3 - \frac{B+C+1}{4}X^2 + \frac{BC}{16}X.$$

This is the minimal Weierstrass model of E. Calculating the discriminant according to the definition, we have

$$\Delta = \frac{(ABC)^2}{2^8}.$$

E has multiplicative reduction at primes dividing Δ. In particular, E is a semi-stable elliptic curve (see §12.1(b)).

Let p be an odd prime. Consider the representation arising from the group of p-torsion points $E[p]$

$$\rho_{E[p]} : G_{\mathbb{Q}} \to GL_2(\mathbb{Z}/p\mathbb{Z}).$$

Denote by K_p the field corresponding to the kernel of this homomorphism. Consider the map

$$\begin{aligned}\det(\rho_{E[p]}) : G_{\mathbb{Q}} &\to (\mathbb{Z}/p/Z)^\times\\ \sigma &\mapsto \det(\rho_p(\sigma)).\end{aligned}$$

Since we have $\det \rho(\mathrm{Frob}_l) = l$ by Theorem 12.4, $\det(\rho)$ is nothing but the Teichmüller character ω (§10.1(e)). That is, if μ_p is the group of all p-th roots of unity in $\overline{\mathbb{Q}}$, then ω is the character satisfying $\sigma(\zeta) = \zeta^{\omega(\sigma)}$ for any $\zeta \in \mu_p$. Note, in particular, that $\mathbb{Q}(\mu_p) \subset K_p$.

Theorem 12.7. *Let l be an odd prime number.*

(i) *Suppose l does not divide ABC and is different from p. Then, l is unramified in K_p.*

(ii) *Suppose l divides ABC and is different from p. Then*

$$l \text{ is unramified in } K_p \Longleftrightarrow \operatorname{ord}_l(ABC) \equiv 0 \mod p.$$

(iii) *If $\operatorname{ord}_p(ABC) \equiv 0 \bmod p$, then p is "moderately ramified." Here, in our situation, "moderately ramified" means the following. If v is a place of K_p lying above p, the completion $(K_p)_v$ of K_p at v can be written in the form*

$$(K_p)_v = k(\mu_p)(\sqrt[p]{u}), \quad u \in \mathcal{O}_{k(\mu_p)}^{\times},$$
$$k \text{ is a finite unramified extension of } \mathbb{Q}_p.$$

To prove Theorem 12.7, we need the theory of Tate curves.

Frey discovered that Fermat's Last Theorem is related to the arithmetic of elliptic curves in the following way. Let p be a prime number greater than or equal to 5. Suppose there exist positive integers a, b, c satisfying

$$a^p + b^p = c^p.$$

Without loss of generality, we may assume that a, b, c are relatively prime, a is odd and b is even. If $a \equiv 3 \bmod 4$, let $A = a^p$, $B = b^p$, $C = c^p$, and if $a \equiv 1 \bmod 4$, let $A = -c^p$, $B = b^p$, $C = -a^p$. Then, consider the elliptic curve

$$E_{(a,b,c)} : y^2 = x(x - B)(x - C).$$

This is called the *Frey curve*. Frey proved, in 1986, that the existence of such a curve contradicts Serre's ε conjecture and the Taniyama-Shimura-Weil conjecture. In other words, Fermat's Last Theorem follows from these two conjectures. This fact granted Fermat's Last Theorem the status of those conjectures that are widely believed to be true. In 1989, Ribet proved Serre's ε conjecture, and thus Fermat's Last Theorem now follows from the Taniyama-Shimura-Weil conjecture. We describe Ribet's theorem in (b).

(b) Ribet's theorem. Let $f = \sum a_n q^n$ be a cusp form of weight $k \geq 2$ and level N that is an eigenfunction of all Hecke operators. If we put $K = \mathbb{Q}_p(\{a_n \mid n \geq 2\})$, then $K/\mathbb{Q}_p$ is a finite extension, and there exists a continuous representation

$$\rho_f : G_{\mathbb{Q}} = \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to GL_2(K)$$

satisfying $\mathrm{Tr}(\rho_f(\mathrm{Frob}_l)) = a_l$ and $\det(\rho_f(\mathrm{Frob}_l)) = l^{k-1}$. This fact was proved by Eichler and Shimura for $k = 2$, as we mentioned in §12.1(f), and by Deligne for the general case.

Since $G_{\mathbb{Q}}$ is compact, we may take ρ_f such that its image is contained in $GL_2(\mathcal{O}_K)$, where $\mathcal{O}_K$ is the integer ring of K. Let π be a generator of the maximal ideal of $\mathcal{O}_K$ and $\mathbb{F}$ its residue field. Consider

$$\rho_f \bmod \pi : G_{\mathbb{Q}} = \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to GL_2(\mathbb{F}).$$

By definition we have

$$\begin{aligned} \mathrm{Tr}(\rho_f \bmod \pi(\mathrm{Frob}_l)) &= a_l \bmod \pi, \\ \det(\rho_f \bmod \pi(\mathrm{Frob}_l)) &= l^{k-1} \bmod \pi. \end{aligned}$$

Thus, if $\rho_f \bmod \pi$ is irreducible, the isomorphism class of $\rho_f \bmod \pi$ is determined uniquely by f. A representation of degree 2 over $\mathbb{F}$ obtained in this way is called modular of weight k and level N.

Serre conjectured that, for a finite field $\mathbb{F}$, an absolutely irreducible representation $\rho_0 : G_{\mathbb{Q}} \to GL_2(\mathbb{F})$ satisfying $\det \rho_0(c) = -1$, where c is the complex conjugation, is always a modular representation. He also conjectured how the weight k and level N are determined by ρ_0 (*Serre's conjecture*). Ribet, following the work of Mazur, solved a part of the Serre conjecture that is called the ε conjecture.

Theorem 12.8 (Mazur, Ribet). *Let $\mathbb{F}$ be a finite field of characteristic $p \geq 5$. Suppose an irreducible representation $\rho_0 : G_{\mathbb{Q}} \to GL_2(\mathbb{F})$ is a modular representation of weight 2 and level N, and N is square free. Let K_{ρ_0} be the extension field corresponding to the kernel of ρ_0, and let l be an odd prime dividing N.*

(i) *If $l \neq p$ and l is unramified in K_{ρ_0}, then ρ_0 is a modular representation of weight 2 and level N/l.*

(ii) *If p divides N and p is "moderately" ramified in K_{ρ_0}, then ρ_0 is a modular representation of weight 2 and lever N/p.*

Let p be a prime number greater than or equal to 5. Suppose that Fermat's Last Theorem is false, and there exits an integer solution

$$a^p + b^p = c^p.$$

Consider the Frey curve $E = E_{(a,b,c)}$ described in (a), and suppose that the Taniyama-Shimura-Weil conjecture is true for E.

Let us apply Theorem 12.8 to the representation $\rho_{E[p]}$ obtained from the p-torsion points. First, we see this representation is irreducible. To do so, we need to use Mazur's theorem on $G_{\mathbb{Q}}$-subgroups of p-torsion points. Then, by the assumption, $E = E_{(a,b,c)}$ is a modular elliptic curve, and thus we see that $\rho_{E[p]}$ is a modular representation of weight 2 and level $N = \prod_{l|\Delta} l$. Then, by Theorems 12.7 and 12.8, $\rho_{E[p]}$ must be a modular representation of weight 2 and level 2. However, there is no cusp form of weight 2 and level 2. This is a contradiction. Therefore, the initial assumption, that is, the existence of a, b and c must be false, and thus Fermat's Last Theorem is proved.

Consequently, to prove Fermat's Last Theorem, it is sufficient to prove that $E_{(a,b,c)}$ is modular. Recall that $E_{(a,b,c)}$ is a semi-stable elliptic curve. Wiles proved the following in 1995.

THEOREM 12.9 (Wiles). *A semi-stable elliptic curve defined over* $\mathbb{Q}$ *is modular.*

COROLLARY 12.10 (Fermat's Last Theorem). *For an integer n greater than or equal to* 3*, there are no positive integers a, b, c satisfying*

$$a^n + b^n = c^n.$$

In the case of $n = 3$ and $n = 4$, Corollary 12.10 had been proved by Euler (and probably by Fermat himself). Thus, we may assume n is a prime number greater than or equal to 5. Hence, Corollary 12.10 follows from the above argument.

In the following we explain the idea of the proof of Theorem 12.9.

(c) Lift of modular Galois representations. Let $\mathcal{O}$ be the integer ring of a local field. If a representation $\rho : G_{\mathbb{Q}} \to GL_2(\mathcal{O})$ is isomorphic to the representation attached to a modular form f, that is, if the corresponding $G_{\mathbb{Q}}$-modules are isomorphic, we say that ρ is coming from a modular form. Then, using what we have said and Faltings' isogeney theorem, we have the following.

THEOREM 12.11. *Let E be an elliptic curve defined over* $\mathbb{Q}$*. The following are equivalent.*

(1) *E is modular.*
(2) *There exists a prime number p such that the representation ρ_p arising from the Tate module $T_p(E)$ is coming from a modular form.*

This is a Galois representation version of the Taniyama-Shimura-Weil conjecture (Conjecture 12.6).

To deduce the fact that an elliptic curve is modular from Theorem 12.11, it suffices to show that ρ_p comes from a modular form for one prime number p. To prove Theorem 12.9, Wiles used mainly the prime number $p = 3$. This is because Langlands and Tunnell had proved that $\rho_{E[3]}$ is modular if $\rho_{E[3]}$ is irreducible (see §11.4), and thus there is a foundation for further arguments.

What Wiles proved precisely was the following theorem. Let p be an odd prime.

THEOREM 12.12 (Wiles). *Let $\mathbb{F}$ be a finite field of characteristic p. Suppose*

$$\rho_0 : G_{\mathbb{Q}} \to GL_2(\mathbb{F})$$

is a modular representation satisfying the following three conditions.

(1) *The restriction of ρ_0 to $\mathrm{Gal}\big(\overline{\mathbb{Q}}/\mathbb{Q}(\sqrt{(-1)^{(p-1)/2}p})\big)$ is absolutely irreducible.*
(2) $\det(\rho_0) = \omega$, *where ω is the Teichmüller character.*
(3) *Some conditions for ρ_0 when it is restricted to the decomposition groups.* (*We omit them here. The condition for ρ_0 when it is restricted to the decomposition group at p is particularly important. If E is a semi-stable elliptic curve, then $\rho_{E[p]}$ satisfies this condition.*)

Let $\mathcal{O}$ be the integer ring of a local field, π a generator of its maximal ideal, and

$$\rho : G_{\mathbb{Q}} \to GL_2(\mathcal{O})$$

a representation satisfying $\rho \bmod \pi = \rho_0$. Furthermore, ρ satisfies the following three conditions.

(i) *If K_ρ is the field corresponding to the kernel of ρ, then the number of primes ramifying in $K_\rho/\mathbb{Q}$ is finite.*
(ii) $\det(\rho) = \kappa$, *where κ is the cyclotomic character.*
(iii) *Some conditions for ρ when it is restricted to the decomposition groups.*

Then, ρ comes from a modular form.

Wiles dealt with the case where $\det(\rho)$ is more general. Here, however, we restrict ourselves to the above case, which we need for the proof of Theorem 12.9.

As we have already mentioned, Langlands and Tunnell proved, in 1981, that an irreducible representation

$$G_{\mathbb{Q}} \to GL_2(\mathbb{F}_3)$$

is modular. For its proof, it is essential that $GL_2(\mathbb{F}_3)$ is solvable. Thus, this proof cannot be generalized to the case of $GL_2(\mathbb{F}_p)$ with $p \geq 5$.

From Theorem 12.2 we see the following immediately.

> "If E is a semi-stable elliptic curve such that the representation $\rho_{E[3]}$ arising from its 3-torsion points is absolutely irreducible even when it is restricted to the absolute Galois group $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}(\sqrt{-3}))$ of $\mathbb{Q}(\sqrt{-3})$, then E is a modular elliptic curve."

Wiles used 5-torsion points as an auxiliary prime in an ingenious manner and successfully removed the irreducibility condition in the above statement. (Here we cannot describe this technique any further.) Finally, he successfully proved that any semi-stable elliptic curve is modular.

(d) $\boldsymbol{R = T}$. Since Theorem 12.12 is the most important theorem in Wiles' proof of Fermat's Last Theorem, we describe its outline here. Theorem 12.12 claims that all the representations that satisfy certain conditions come from modular forms. Note that it resembles the method of the proof of the Iwasawa main conjecture by Mazur and Wiles (§10.3(e)), which "constructs all the unramified extensions from modular forms." Mazur's deformation theory of Galois representations plays an essential role in the proof.

Let ρ_0 be as in Theorem 12.12. For representations ρ satisfying the conditions (i), (ii) and (iii) in Theorem 12.12, there exist a ring R and a Galois representation

$$\rho_R : G_{\mathbb{Q}} \to GL_2(R)$$

satisfying the conditions below. For ρ satisfying (i), (ii) and (iii) in Theorem 12.12, there is a ring homomorphism

$$R \to \mathcal{O}$$

such that the composition

$$G_{\mathbb{Q}} \xrightarrow{\rho_R} GL_2(R) \to GL_2(O)$$

of ρ_R and the homomorphism $GL_2(R) \to GL_2(\mathcal{O})$ induced by this ring homomorphism is isomorphic to ρ. On the other hand, if we add the condition "ρ comes from a modular form" to (i), (ii) and (iii), there still exist a universal ring T and a Galois representation

$$\rho_T : G_{\mathbb{Q}} \to GL_2(T)$$

such that any modular representation ρ satisfying these conditions is obtained from ρ_T. In other words, if there is such a ρ, then we have a ring homomorphism $T \to O$ such that the composition

$$G_{\mathbb{Q}} \xrightarrow{\rho_T} GL_2(T) \to GL_2(\mathcal{O})$$

of ρ_T and the homomorphism $GL_2(T) \to GL_2(\mathcal{O})$ induced by this ring homomorphism is isomorphic to ρ. Using the language of deformation theory, Theorem 12.12 can be paraphrased simply by

$$R = T.$$

By the definitions of R and T, there is a ring homomorphism

$$R \to T,$$

and it is surjective. Our aim is to prove this map is an isomorphism.

The following theorem is proved by H. W. Lenstra Jr., who made improvements to Wiles' method. Since Lenstra's theorem is easier to state than Wiles' original theorem, we state it here.

Theorem 12.13 (Lenstra Jr.). *Let $\mathcal{O}$ be the ring of integers of a local field, R a complete Noetherian local ring over $\mathcal{O}$, and T a finite flat Noetherian local ring over $\mathcal{O}$. Suppose there exist homomorphisms of $\mathcal{O}$-algebras $\varphi : R \to T$ and $\psi : T \to \mathcal{O}$. Let $\phi = \psi \circ \varphi$, and let* $\mathrm{Ann}_T(\mathrm{Ker}\,\psi)$ *be the annihilator of* $\mathrm{Ker}\,\psi$. *If the inequality*

$$\#((\mathrm{Ker}\,\phi)/(\mathrm{Ker}\,\phi)^2) \leq \#(\mathcal{O}/\varphi(\mathrm{Ann}_T(\mathrm{Ker}\,\psi))),$$

holds, then $\psi : R \xrightarrow{\simeq} T$ is an isomorphism.

Thereby the problem is reduced to examining $((\mathrm{Ker}\,\phi)/(\mathrm{Ker}\,\phi)^2)$. This leads us to study the Selmer group of $\mathrm{Sym}^2 f$ of a modular form f. The inequality in Theorem 12.13 is in fact an equality, and it is equivalent to the formula

$$\begin{aligned}&\#(\text{Selmer group of } \mathrm{Sym}^2 f)\\ &\qquad = (p \text{ component of the algebraic part of } L(2, \mathrm{Sym}^2 f)).\end{aligned}$$

$\operatorname{Ker}\phi/(\operatorname{Ker}\phi)^2$ in Theorem 12.13 corresponds to the left-hand side in the above formula, and $\mathcal{O}/\psi(\operatorname{Ann}_T(\operatorname{Ker}\psi))$ corresponds to the right-hand side. (For $L(s, \operatorname{Sym}_2 f)$, see §9.1(e).) In this way a ζ function (L-function) emerges again and plays an important role. It is profoundly mysterious that ζ functions always emerge in any important occasions. In the case of ideal class groups, the above equality about Selmer groups corresponds to

$$\# A_{\mathbb{Q}(\mu_p)}^{\omega^i} = \#(\mathbb{Z}_p/L(0, \omega^{-i})\mathbb{Z}_p),$$

where i is an odd integer satisfying $1 < i < p-1$. (See Theorem 10.37 in §10.3(c).) In this way, one may think that the equality in question is something Iwasawa theory can handle.

While the strategy of the proof of the Iwasawa main conjecture is to construct all unramified extensions from modular forms, the strategy here is also to construct all appropriate representations from modular forms, and they are comparable. In order to prove the Iwasawa main conjecture, we need a representation to GL_2 in order to construct abelian extensions. In terms of §9.4(b), we need to view the ground floor from upstairs. Here, in order to study 2-dimensional representations, we need to view the second floor from the third floor.

Unfortunately, we do not have enough space here to describe the proof of the inequality concerning Selmer groups, but we will remark that Wiles proved this inequality when T is a complete intersection. The notion of a complete intersection is related to the singularities of local rings, and it means that the situation of singularities is not too bad. This last hurdle was cleared by the paper Wiles coauthored with R. Taylor.

THEOREM 12.14 (Taylor, Wiles). *T is a complete intersection.*

The proof shrewdly uses the Hecke algebra of modular forms of level $Nl_1 \cdots l_r$.

Summary

12.1. In the arithmetic of elliptic curves defined over the rational number field, we encounter the same sequence of integers $(a_l)_{l \in \mathcal{L}}$ when we count the number of rational points over finite fields, when we study the action of the Galois group on the Tate modules, or when we study the ζ functions. The Taniyama-Shimura-Weil conjecture states that we can construct a modular form from this important sequence $(a_l)_{l\in\mathcal{L}}$.

12.2. Wiles proved the Taniyama-Shimura-Weil conjecture for semi-stable elliptic curves over the rational number field. As a consequence, Fermat's Last Theorem was proved.