Show that the diamond operator in characteristic $p$ on the moduli space and on the modular curve (cf. diagram (8.29)) are compatible in that the following diagram commutes:

\[
\begin{array}{ccc}
\text{Div}^0(\tilde{S}_1(N)') & \xrightarrow{\langle d \rangle} & \text{Div}^0(\tilde{S}_1(N)) \\
\downarrow & & \downarrow \\
\text{Pic}^0(\tilde{X}_1(N)) & \xrightarrow{\langle d \rangle^*} & \text{Pic}^0(\tilde{X}_1(N)) \\
\end{array}
\]

(8.39)

(d) Use part (b) and diagram (8.39) to complete the proof that diagram (8.34) commutes.

### 8.8 Fourier coefficients, $L$-functions, and Modularity

Recall that if $E$ is an elliptic curve over $\mathbb{Q}$ then its analytic conductor was defined in Section 7.7 as the smallest $N$ such that $X_0(N)$ surjects to $E$, and its algebraic conductor $N_E$ was described in Section 8.3. Both conductors depend only on the isogeny class over $\mathbb{Q}$ of $E$. In this section we simply call the algebraic conductor the conductor en route to explaining why the two conductors are in fact equal and no longer need to be distinguished.

**Theorem 8.8.1 (Modularity Theorem, Version $a_p$).** Let $E$ be an elliptic curve over $\mathbb{Q}$ with conductor $N_E$. Then for some newform $f \in S_2(\Gamma_0(N_E))$, $a_p(f) = a_p(E)$ for all primes $p$.

This version of Modularity is most obviously related to Version $A_Q$, the version that provides a map $A'_f \to E$, since each version involves a newform $f$, and unsurprisingly the two $f$’s are the same. But since $A'_f$ is a variety rather than a curve, and our policy is to argue using only curves, we give instead a partial proof that Version $X_Q$, providing a map $X_0(N) \to E$, implies this version. The argument necessarily requires a little effort to extract $f$ from $X_0(N)$ in consequence of avoiding varieties. Specifically, we prove

**Theorem 8.8.2.** Let $E$ be an elliptic curve over $\mathbb{Q}$ with conductor $N_E$, let $N$ be a positive integer, and let

$$\alpha : X_0(N) \to E$$
be a nonzero morphism over \( \mathbb{Q} \) of curves over \( \mathbb{Q} \). Then for some newform \( f \in S_2(\Gamma_0(M_f)) \) where \( M_f \mid N \),

\[
ap_p(f) = a_p(E) \quad \text{for all primes } p \nmid N E N.
\]

After proving this we will touch on the rest of the argument that Version \( X_\mathbb{Q} \) of Modularity implies Version \( a_p \), and on the more natural-seeming argument starting from Version \( A_\mathbb{Q} \).

**Proof.** The route from \( a_p(f) \) for some \( f \) to \( a_p(E) \) is that

- \( a_p(f) \) on \( A'_f \) is \( T_p \) for each \( f \), by a variant of diagram (6.15), and a sum of factors \( A'_f \) over all \( f \) is isogenous to \( \text{Pic}^0(X_0(N)) \), by a variant of Theorem 6.6.6,
- \( T_p \) on \( \text{Pic}^0(X_0(N)) \) reduces to \( \sigma_{p,*} + \sigma_{p,*}^* \) on \( \text{Pic}^0(\tilde{X}_0(N)) \), by the Eichler–Shimura Relation \( (X_0(N)) \) has good reduction at \( p \) because \( p \nmid N \),
- \( \sigma_{p,*} + \sigma_{p,*}^* \) on \( \text{Pic}^0(\tilde{X}_0(N)) \) commutes with \( \tilde{\alpha}_* \) to become \( \sigma_{p,*} + \sigma_{p,*}^* \) on \( \text{Pic}^0(\tilde{E}) \), by formulas (8.17) and (8.19) \( (E) \) has good reduction at \( p \) because \( p \nmid N E \),
- and finally \( \sigma_{p,*} + \sigma_{p,*}^* \) on \( \text{Pic}^0(\tilde{E}) \) is \( a_p(E) \), by Proposition 8.3.2.

Recall some ideas from Section 6.6, given there for the group \( \Gamma_1(N) \) and now modified for \( \Gamma_0(N) \). The complex vector space \( S_2(\Gamma_0(N)) \) has basis

\[
B'_2(N) = \bigcup_f \bigcup_n \bigcup_{\sigma} f^\sigma(n \tau)
\]

where the first union is taken over equivalence class representatives of newforms \( f \in S_2(\Gamma_0(M_f)) \) with \( M_f \) dividing \( N \), the second over divisors \( n \) of \( N/M_f \), and the third over embeddings \( \sigma : K_f \hookrightarrow \mathbb{C} \). Work over \( \mathbb{C} \) now, identifying complex algebraic curves and Riemann surfaces, and identifying Picard groups and Jacobians. Then the Picard group associated to \( \Gamma_0(N) \) is isogenous to a direct sum of Abelian varieties, both sides being viewed as complex tori,

\[
\text{Pic}^0(X_0(N)\mathbb{C}) \twoheadrightarrow \bigoplus_{f,n} A'_{f,C},
\]

and there exists a dual isogeny

\[
\bigoplus_{f,n} A'_{f,C} \rightarrow \text{Pic}^0(X_0(N)\mathbb{C}).
\]

The given map \( \alpha : X_0(N) \rightarrow E \) extends to

\[
\alpha_C : X_0(N)\mathbb{C} \rightarrow E\mathbb{C},
\]

viewed as a morphism of complex algebraic curves or as a holomorphic map of compact Riemann surfaces, a surjection in either case.

For any \( p \nmid N E N \) the diagram
If \( a_p(f) \neq a_p(E) \) then the top row surjects from \( \bigoplus_n A'_{f,C} \) to \( \bigoplus_n A'_{f,C} \).

(b) The square commutes.

(c) The composite map on the bottom row is zero.

If \( a_p(f) \neq a_p(E) \) for some \( f \) and some \( p \nmid N_EN \) then in the diagram for that \( p \), mapping \( \bigoplus_n A'_{f,C} \) across the top row takes it to all of \( \bigoplus_n A'_{f,C} \) by property (a), and then mapping this down gives its isogenous image in \( \text{Pic}^0(X_0(N)|C) \). By property (b) the isogenous image also comes from mapping \( \bigoplus_n A'_{f,C} \) down the left side of the diagram and then halfway across the bottom row. Property (c) now shows that the isogenous image of \( \bigoplus_n A'_{f,C} \) lies in \( \ker(\alpha_{C,s}) \). Therefore, if for each \( f \) there exists a \( p \mid N_EN \) such that \( a_p(f) \neq a_p(E) \) then all of \( \text{Pic}^0(X_0(N)|C) \) lies in \( \ker(\alpha_{C,s}) \). But \( \alpha_{C,s} \) surjects, so this is impossible. That is, there is a newform \( f \) such that \( a_p(f) = a_p(E) \) for all \( p \nmid N_EN \), as we needed to prove.

Returning to diagram (8.40), to prove its property (a) let \( a_p(f) \neq a_p(E) \). Recall that \( a_p(f) \) is an algebraic integer, and \( a_p(E) \) is a rational integer. Thus their difference \( \delta \) satisfies a minimal monic polynomial with rational integer coefficients,

\[
\delta^c + a_1\delta^{c-1} + \cdots + a_{c-1}\delta + a_c = 0, \quad a_1, \ldots, a_c \in \mathbb{Z}, \quad a_c \neq 0.
\]

The resulting relation \( \delta(\delta^{c-1} + a_1\delta^{c-2} + \cdots + a_{c-1}) = -a_c \) shows that \( \delta \) is surjective on \( A'_{f,C} \) as desired, since \(-a_c \) is.

Property (b) follows quickly from diagram (6.20) suitably modified from \( \Gamma_1(N) \) to \( \Gamma_0(N) \).

To prove property (c) switch back to working over \( \mathbb{Q} \) and consider the following diagram:

\[
\begin{array}{cccc}
\text{Pic}^0(X_0(N)) & \xrightarrow{T_p-a_p(E)} & \text{Pic}^0(X_0(N)) & \xrightarrow{\alpha_{\ast}} & \text{Pic}^0(E) \\
\downarrow & & \downarrow & & \downarrow \\
\text{Pic}^0(\bar{X}_0(N)) & \xrightarrow{\sigma_p \ast + \sigma_p \ast - a_p(E)} & \text{Pic}^0(\bar{X}_0(N)) & \xrightarrow{\tilde{\alpha}_{\ast}} & \text{Pic}^0(\bar{E}) \\
\downarrow & \quad 1 & \downarrow & \quad 1 & \downarrow \\
\text{Pic}^0(\bar{X}_0(N)) & \xrightarrow{\tilde{\alpha}_{\ast}} & \text{Pic}^0(\bar{E}) & \xrightarrow{\sigma_p \ast + \sigma_p \ast - a_p(E)} & \text{Pic}^0(\bar{E}).
\end{array}
\]

The top row here is a restriction of the bottom row of diagram (8.40), and it suffices to prove that this restriction doesn’t surject (Exercise 8.8.1). The top
left square commutes since it is essentially (8.38) from the Eichler–Shimura Relation, and the top right square commutes by Corollary 8.5.9. The bottom rectangle commutes because formulas (8.17) and (8.19) from earlier in the chapter show that \( \sigma_p \) and \( \sigma_p^* \) commute with \( \tilde{\alpha}_* \), and multiplication by \( a_p(E) \) certainly commutes with the linear map \( \tilde{\alpha}_* \) as well. The second map on the bottom row is zero by Proposition 8.3.2, making the bottom row zero and thus making the middle row zero. So the top row followed by the right vertical arrow to the second row is zero. But the vertical arrow surjects by Theorem 8.5.9. This shows that the top row can’t surject, as desired. With properties (a), (b), and (c) of diagram (8.40) established, the proof of Theorem 8.8.2 is complete. \( \square \)

For Version \( X_Q \) of Modularity to imply Version \( a_p \), the newform \( f \) of Theorem 8.8.2 needs to have level \( M_f = N_E \) and to have Fourier coefficients \( a_p(f) = a_p(E) \) for all \( p \).

Strong Multiplicity One shows that all the Fourier coefficients of \( f \) are rational integers. To see this, recall that if \( \sigma \) is any automorphism of \( Q \) then the conjugate of \( f \) under \( \sigma \) is defined as

\[
 f^\sigma(\tau) = \sum_{n=1}^{\infty} a_n(f)^\sigma q^n.
\]

According to Theorem 6.5.4 all conjugates of \( f \) are again newforms. In this case they are all equal to \( f \) itself because for all \( p \nmid N \) and all \( \sigma \),

\[ a_p(f) = a_p(E) \implies a_p(f) \in \mathbb{Z} \implies a_p(f)^\sigma = a_p(f), \]

so that \( f^\sigma = f \) by Strong Multiplicity One. This means that the Fourier coefficients are algebraic integers invariant under \( \text{Aut}(Q) \), i.e., the Fourier coefficients are rational integers as claimed. Thus \( f \) has number field \( K_f = \mathbb{Q} \).

Obtaining Version \( a_p \) of Modularity from here requires results beyond the scope of this book, but we sketch the ideas. As quoted in Section 7.7, the Abelian variety \( A_f' \), viewed as a variety over \( Q \), has dimension \( [K_f : Q] = 1 \), i.e., \( A_f' \) is an elliptic curve over \( Q \). The map \( X_0(N) \to \text{Pic}^0(X_0(N)) \to A_f' \) is defined over \( Q \), so we can run the proof of Theorem 8.8.2 again to show that there exists a newform \( g \) such that \( a_p(g) = a_p(A_f') \) for all but finitely many \( p \). The proof in this case shows that \( g = f \), so

\[ a_p(f) = a_p(A_f') \quad \text{for all but finitely many } p. \] (8.42)

Carayol [Car86], building on the work of Eichler–Shimura, Langlands, and Deligne, showed that in fact \( a_p(f) = a_p(A_f') \) for all \( p \), and the level of \( f \) is the conductor of \( A_f' \), notated \( M_f = N_f \), and this is also the analytic conductor of \( A_f \). But also, the work so far gives an isogeny over \( Q \) from \( A_f' \) to \( E \), so that \( a_p(A_f') = a_p(E) \) for all \( p \) and the conductor of \( A_f' \) is the conductor of \( E \), that is, \( N_f = N_E \). This gives \( a_p(f) = a_p(E) \) for all \( p \) and \( M_f = N_E \), i.e., Version \( a_p \).
of Modularity. It also shows that the analytic and algebraic conductors of \(E\) are the same. That is, the newform \(f\) associated to \(E\) has the same level as the smallest modular curve \(X_0(N)\) that maps to \(E\).

Starting instead from Version \(A_Q\) of Modularity, that is, starting from a map \(\alpha : A'_f \rightarrow E\), a more natural-looking proof that \(a_p(f) = a_p(E)\) for all \(p \nmid N_E M_f\) (where \(N_E\) is the conductor of \(E\) and \(M_f\) is the level of \(f\)) would set up a diagram analogous to (8.41),

\[
\begin{array}{ccc}
A'_f & \xrightarrow{a_p(f) - a_p(E)} & A'_f \\
\downarrow & & \downarrow \\
\tilde{A'}_f & \xrightarrow{\sigma_p + \sigma^*_p - a_p(E)} & \tilde{A'}_f \\
\downarrow & & \downarrow \\
A'_f & \xrightarrow{\tilde{\alpha}} & \tilde{A'}_f \\
\downarrow & & \downarrow \\
E & \xrightarrow{\sigma_p + \sigma^*_p - a_p(E)} & \tilde{E} \\
\end{array}
\]

If \(a_p(f) \neq a_p(E)\) then the top row is surjective, as is \(E \rightarrow \tilde{E}\), but the bottom row is zero and hence the middle row is zero, giving a contradiction because all the rectangles commute. However, this argument makes use of the structure of \(A'_f\) as a variety over \(Q\), reduces the variety modulo \(p\), and then makes further use of the structure of the reduction as a variety over \(F_p\), invoking algebraic geometry well beyond the scope of this book. By contrast our proof of Theorem 8.8.2 makes no reference to the variety structure of \(\text{Pic}^0(X_0(N))\), and it uses \(\text{Pic}^0(\tilde{X}_0(N))\), the Picard group of the reduced curve, rather than reducing an Abelian variety.

Section 1.3 mentioned that some complex tori have complex multiplication, endomorphisms other than \(\{[N] : N \in \mathbb{Z}\}\). This notion extends to algebraic elliptic curves, providing more examples of Modularity. For example the elliptic curve \(E : y^2 = x^3 - d, \quad d \in \mathbb{Z}, \quad d \neq 0\)

has the order 3 automorphism \((x, y) \mapsto (\mu_3 x, y)\) over \(\mathbb{Q}(\mu_3)\), and its ring of endomorphisms is isomorphic to \(A = \mathbb{Z}[\mu_3]\). As in Exercise 8.3.6(b), \(a_p(E) = 0\) for all \(p \equiv 2 \pmod{3}\) for which the displayed Weierstrass equation is minimal. The theory of complex multiplication in fact describes \(a_p(E)\) for all \(p\): there exist an integer \(N \mid 12 \prod_{p|d} p\) and a character \(\chi : (A/NA)^* \rightarrow \mathbb{C}^*\) of order 6 such that \(\chi(u) = u^{-1}\) for \(u \in A^*, \chi(a) = (a/3)\) (Legendre symbol) for \(a \in (\mathbb{Z}/N\mathbb{Z})^*\), and

\[
a_p(E) = \frac{1}{6} \sum_{\substack{n \in A \\ |n|^2 = p}} \chi(n)n.
\]

(See Chapter 2 of [Sil94] for proofs of these results, and see Exercise II.5.7 of [Kob93] for the case \(d = -16\), giving \(N = 3\).) The solution-counts of \(E\) are
the Fourier coefficients of the(338,65),(591,207)

\[ \theta_{2,\chi}(\tau) = \frac{1}{\pi} \sum_{n \in A} \chi(n) ne^{2\pi i |n|^2 \tau}, \quad \tau \in \mathcal{H}. \]

This function is a modular form of level $3N^2$ by the weight 2 version of Hecke’s construction in Section 4.11. It is a normalized eigenform by an argument similar to the one in Section 5.9 for the weight 1 theta function. For the minimal choice of $N$, it is in fact a newform of level $3N^2 = N_E$, thus illustrating Version $a_p$ of Modularity for these elliptic curves $E$. The elliptic curves $E$ over $\mathbb{Q}$ with complex multiplication form a small class of examples—for example they have only a finite set of $j$-invariants—but they are important in number theory, and Shimura’s proof of Version $X \mathbb{Q}$ of Modularity for such curves [Shi71] provided evidence for the algebraic formulations of Modularity in general.

Version $a_p$ of the Modularity Theorem rephrases in terms of $L$-functions. Recall from Chapter 5 that if $f \in \mathcal{S}_2(\Gamma_0(N))$ is a newform then its $L$-function is

\[ L(s, f) = \sum_{n=1}^{\infty} a_n(f)n^{-s} = \prod_p (1 - a_p(f)p^{-s} + 1_N(p)p^{1-2s})^{-1}, \quad (8.43) \]

with convergence in a right half plane. For any elliptic curve $E$ over $\mathbb{Q}$ let $1_E$ be the trivial character modulo the conductor $N$ of $E$ as in Section 8.3. The Hasse–Weil $L$-function of $E$, encoding all the solution-counts $a_p(E)$, is

\[ L(s, E) = \prod_p (1 - a_p(E)p^{-s} + 1_E(p)p^{1-2s})^{-1}. \]

From Section 8.3 we have the definitions

\[ a_p(E) = 1, \quad a_{pe}(E) = p^e + 1 - |\overline{E}(\mathbb{F}_p^e)|, \quad e \geq 1 \]

and the recurrence

\[ a_{pe}(E) = a_p(E)a_{pe-1}(E) - 1_E(p)pa_{pe-2}(E), \quad e \geq 2. \]

One more definition,

\[ a_{mn}(E) = a_m(E)a_n(E), \quad (m, n) = 1, \]

completes the analogy between the values $a_n(E)$ and the Fourier coefficients $a_n(f)$ of a newform. As in Chapter 5, the $L$-function of an elliptic curve is now

\[ L(s, E) = \sum_{n=1}^{\infty} a_n(E)n^{-s} = \prod_p (1 - a_p(E)p^{-s} + 1_E(p)p^{1-2s})^{-1}. \quad (8.44) \]

Half plane convergence of $L(s, E)$ can be established as well by estimating $a_p(E)$. But this is unnecessary since comparing the products in (8.43) and (8.44) shows that Version $a_p$ of Modularity is equivalent to
Theorem 8.8.3 (Modularity Theorem, Version L). Let $E$ be an elliptic curve over $\mathbb{Q}$ with conductor $N_E$. Then for some newform $f \in S_2(\Gamma_0(N_E))$,

$$L(s, f) = L(s, E).$$

Faltings’s Isogeny Theorem (Corollary 5.2 in Chapter 2 of [CS86]), a deep result, now shows that this Version $L$ of Modularity implies

Theorem 8.8.4 (Modularity Theorem, strong Version $A_\mathbb{Q}$). Let $E$ be an elliptic curve over $\mathbb{Q}$ with conductor $N_E$. Then for some newform $f \in S_2(\Gamma_0(N_E))$ the Abelian variety $A'_f$ is also an elliptic curve over $\mathbb{Q}$ and there exists an isogeny over $\mathbb{Q}$

$$A'_f \rightarrow E.$$

To see this, suppose that by Version $L$ we have $L(s, f) = L(s, E)$. Then $f$ has rational coefficients, making $A'_f$ an elliptic curve. Equation (8.42) shows that $L(s, A'_f)$ and $L(s, f)$ have the same Euler product factors for all but finitely many primes $p$. So the same is true of $L(s, A'_f)$ and $L(s, E)$, and now Faltings’s Theorem gives an isogeny $A'_f \rightarrow E$.

Version $L$ of the Modularity Theorem shows that the half plane convergence, analytic continuation, and functional equation of $L(s, f)$ from Theorem 5.10.2 now apply to $L(s, E)$. This is important because the continued $L(s, E)$ is conjectured to contain sophisticated information about the group structure of $E$. Specifically, since $E(\mathbb{Q})$ is a finitely generated Abelian group it takes the form

$$E(\mathbb{Q}) \cong T \oplus \mathbb{Z}^r,$$

where $T$ is the torsion subgroup and $r$ is the rank. The rank is much harder to compute than the torsion. However,

Conjecture 8.8.5 (Weak Birch and Swinnerton-Dyer Conjecture). Let $E$ be an elliptic curve defined over $\mathbb{Q}$. Then the order of vanishing of $L(s, E)$ at $s = 1$ is the rank of $E(\mathbb{Q})$. That is, if $E(\mathbb{Q})$ has rank $r$ then

$$L(s, E) = (s - 1)^r g(s), \quad g(1) \neq 0, \infty.$$

The original half plane of convergence of $L(s, E)$ is $\{\text{Re}(s) > 2\}$, and the functional equation then determines $L(s, E)$ for $\text{Re}(s) < 0$, but the behavior of $L(s, E)$ at the center of the remaining strip $\{0 \leq \text{Re}(s) \leq 2\}$ is what conjecturally determines the rank of $E(\mathbb{Q})$. The Birch and Swinnerton-Dyer Conjecture would give an algorithm for finding all rational points on elliptic curves, and it would give an effective method for finding imaginary quadratic fields with a given class number. For more on the Birch and Swinnerton-Dyer Conjecture see [Tat02], Appendix C.16 of [Sil86], or Chapter 17 of [Hus04].
Exercise

8.8.1. Let $\beta_C$ denote the map of the bottom row of diagram (8.40),

$$\beta_C : \text{Pic}^0(X_0(N)_C) \xrightarrow{\alpha \circ (T_p - a_p(E))} \text{Pic}^0(E_C).$$

Augment $\beta_C$ to get a map of complex algebraic curves,

$$\gamma : X_0(N) \rightarrow \text{Pic}^0(X_0(N)_C) \xrightarrow{\beta_C} \text{Pic}^0(E_C) \rightarrow E.$$

Here the first stage is $P \mapsto [(P) - (P_0)]$ where $P_0 \in X_0(N)_C$ is a base point, and the third stage is $\sum(Q_i) \mapsto \sum Q_i$. Since $\gamma_C$ is a composite of holomorphic maps, it is holomorphic as a map of Riemann surfaces and therefore it is a morphism as a map of complex algebraic curves.

(a) Show that if $\gamma_C$ is zero then $\beta_C$ is zero. (A hint for this exercise is at the end of the book.)

(b) Assume that the base point $P_0$ in part (a) has algebraic coordinates. Show that $\gamma_C$ is defined over $\mathbb{Q}$ as follows. It suffices to show that $\gamma_{C \sigma} = \gamma_C$ for all $\sigma \in \text{Aut}(C/\mathbb{Q})$. Compute that for any $P \in X_0(N)_C$ and any $\sigma$,

$$\gamma_C(P^\sigma) = \gamma_C(P)^\sigma = \gamma_{C \sigma}(P^\sigma).$$

Since $P^\sigma$ can be any point of $X_0(N)_C$, this gives the result.

(c) Let $\beta$ denote the map of the top row of diagram (8.41),

$$\beta : \text{Pic}^0(X_0(N)) \xrightarrow{\alpha \circ (T_p - a_p(E))} \text{Pic}^0(E).$$

This is the restriction of $\beta_C$ to $\mathbb{Q}$-points. Consider the corresponding restriction of $\gamma_C$, viewed as a morphism over $\mathbb{Q}$ of algebraic curves over $\mathbb{Q}$ according to part (b),

$$\gamma : X_0(N) \rightarrow \text{Pic}^0(X_0(N)) \xrightarrow{\beta} \text{Pic}^0(E) \rightarrow E.$$

Use the maps of curves $\gamma$ and $\gamma_C$ to show that if $\beta$ does not surject then $\beta_C$ is zero. This was used in proving property (c) of diagram (8.40). The result is immediate from quoting that as morphisms from varieties to curves, $\beta$ and $\beta_C$ are both constant or surjective, but the argument in this exercise uses only the algebraic geometry of curves.