Vol. 61, No. 2 DUKE MATHEMATICAL JOURNAL © October 1990

A TAMENESS CRITERION FOR GALOIS REPRESENTATIONS ASSOCIATED TO MODULAR FORMS (MOD p)

BENEDICT H. GROSS

We begin by recalling some results on the 2-dimensional Galois representations which are associated to modular forms (mod p). If $f = \Sigma a_n q^n$ is a normalized cuspidal eigenform of weight k and character ε for $\Gamma_1(N)$, with coefficients in a finite field E of characteristic p, there is a continuous semi-simple Galois representation

$$\rho_f\colon \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2(E)$$

which is characterized as follows. The representation ρ_f is unramified for all primes $l \nmid Np$, and the matrix $\rho_f(\mathrm{Frob}_l)$ has characteristic polynomial $x^2 - a_l x + \varepsilon(l)l^{k-1}$. The representation ρ_f was conjectured to exist by Serre (cf. [S2], [S6]) and its existence was proved by Deligne (cf. [D1] for the case $N = 1$, and [C] for more general levels). When $k \geqslant 2$ and $a_p \neq 0$, Deligne [D2] also proved that the restriction of ρ_f to a decomposition group at p in $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ has image contained in a Borel subgroup of $\mathrm{GL}_2(E)$. Up to conjugation, this restriction has the form

$$(0.1) \qquad \begin{pmatrix} \chi^{k-1} \cdot \lambda(\varepsilon(p)/a_p) & * \\ 0 & \lambda(a_p) \end{pmatrix}$$

where χ is the character of $\mathrm{Gal}(\bar{\mathbb{Q}}_p/\mathbb{Q}_p)$ giving its action on μ_p and, for any $\alpha \in E^*$, $\lambda(\alpha)$ is the unramified character taking Frob_p to α.

In this paper, we will establish a modular criterion conjectured by Serre [S7, pg. 18] for the representation ρ_f to be tamely ramified at p, or more precisely, for $* = 0$ in (0.1). Assume that f has weight $2 \leqslant k \leqslant p$ and $a_p \neq 0$; when $k = p$ assume further that $a_p^2 \neq \varepsilon(p)$, so the two characters $\chi^{k-1}\lambda(\varepsilon(p)/a_p)$ and $\lambda(a_p)$ are distinct. The criterion says that ρ_f is completely reducible when restricted to $\mathrm{Gal}(\bar{\mathbb{Q}}_p/\mathbb{Q}_p)$ if and only if there is a normalized eigenform $g = \Sigma b_n q^n$ of weight $k' = p + 1 - k$ and character ε for $\Gamma_1(N)$ over E, whose Fourier coefficients satisfy $n^k b_n = n a_n$ for all $n \geqslant 1$.

The relationship between f and g is symmetric (for example, the relation between Fourier coefficients may be written $nb_n = n^{k'} a_n$), and Serre calls the pair (f, g) of normalized eigenforms "companions". An equivalent formulation of companionship is that the Galois representations ρ_f and ρ_g satisfy: $\rho_f \otimes \chi \simeq \rho_g \otimes \chi^k$. Using this, it is easy to show that the existence of a companion forces ρ_f to be completely

Received October 23, 1989. Revision received February 15, 1990.

reducible when restricted to $\mathrm{Gal}(\bar{\mathbb{Q}}_p/\mathbb{Q}_p)$: one finds that the character $\lambda(a_p)$ occurs as a direct summand in $\rho_f \simeq \rho_g \otimes \chi^{k-1}$. The converse is more difficult, and provides an excellent test case for Serre's general conjectures ([S8], [S9]) on the weight of a "modular" Galois representation. We note that the case when $k = p$ and $a_p^2 = \varepsilon(p)$ is still open, and seems to require new methods.

We begin by recalling the geometric theory of modular forms, due to Deligne-Rapoport [DR] and Katz [K2]. We then consider the case of modular forms (mod p), which is due to Serre and Swinnerton-Dyer [S4], [Sw]. Our approach stresses the Igusa coverings of modular curves in characteristic p, and we prove a theorem of Serre on the differentials of the Igusa curve which frequently allows one to reduce questions on forms of weight $k \leqslant p + 1$ on $\Gamma_1(N)$ (mod p) to forms of weight 2 on $\Gamma_1(Np)$. The construction of the representation ρ_f when $k \leqslant p + 1$ is accomplished using the p-torsion in the Jacobian of the curve $X_1(Np)$; this construction is based on ideas of Fontaine and Serre [F3], [S7] and is well-suited to the study of the local behavior of ρ_f at p. We have also followed Mazur's approach [M], using ideals in the Hecke algebra, fairly closely. The restriction of ρ_f to a decomposition group at p is determined, when $a_p \neq 0$, using the theory of ordinary p-divisible groups.

The proof of Serre's conjecture on companion forms uses p-adic techniques, and specifically the different p-adic cohomology theories (de Rham, crystalline, Washnitzer-Monsky) of modular curves and their Jacobians. Here we confess that we have occasionally used rather artificial methods for defining the action of Hecke operators on these cohomology groups, and have not always checked that the actions are compatible with isomorphisms between the theories. In particular, the assertions preceding (15.4), (15.7), and (16.7) depend on an unchecked compatibility.

It is a pleasure to thank O. Atkin, R. Coleman, N. Elkies, N. Katz, and B. Mazur for their help. Special thanks go to J.-P. Serre, whose beautiful conjectures stimulated my interest in this subject, and who provided invaluable assistance in the writing of this paper.

Table of Contents

§1. Elliptic curves. This section contains a brief review of the theory of elliptic curves, and generalized elliptic curves, over an arbitrary base scheme S. We refer to the papers of Deligne [D3], Deligne-Rapoport [DR], and Katz-Mazur [KM] for the proofs.

An elliptic curve over a field F is a complete, nonsingular curve of genus 1 over F, furnished with an F-rational point. An elliptic curve E over a scheme S is a proper and smooth morphism $\pi\colon E \to S$, furnished with a section $e\colon S \to E$, whose geometric fibres are all elliptic curves. If $S = \operatorname{Spec} R$ is affine, we will often refer to E as an elliptic curve over R. The addition law on the fibres gives E the structure of a commutative group scheme over S [KM, 2.1].

Let $N \geqslant 1$ be an integer, and let E_N be the kernel of multiplication by N on E. Then E_N is a finite flat group scheme of rank N^2 over S; if N is invertible on S, then E_N is an étale group scheme locally isomorphic to $(\mathbb{Z}/N\mathbb{Z})^2$ [KM, 2.3]. There is a canonical, strictly alternating pairing [KM, 2.8]

$$e_N\colon E_N \times E_N \to \mu_N \tag{1.1}$$

where μ_N is the group scheme of N^{th} roots of unity. The e_N pairing is non-degenerate, in the sense that the map

$$\begin{gathered} f_N\colon E_N \to {}^tE_N = \operatorname{Hom}(E_N, \mathbb{G}_m) \\ \alpha \mapsto (\beta \mapsto e_N(\alpha, \beta)) \end{gathered} \tag{1.2}$$

defines an isomorphism between E_N and its Cartier dual ${}^tE_N$.

The invertible sheaf $\Omega^1_{E/S}$ on E has degree zero on each fibre, and the trace map of Serre-Grothendieck duality defines an isomorphism $R^1\pi_*\Omega^1_{E/S} \simeq \mathcal{O}_S$. Hence

$$\underline{\omega}_E \underset{\text{def}}{=} \pi_*\Omega^1_{E/S} \tag{1.3}$$

is an invertible sheaf on S, whose formation commutes with change of base [KM, 2.2]. Its dual is the invertible sheaf $R^1\pi_*\mathcal{O}_{E/S}$, which is isomorphic to the sheaf $\underline{\operatorname{Lie}}(E)$ of Lie algebras on S. Since π is smooth, $\Omega^1_{E/S} = \Omega^{reg}_{E/S}$ and one also has a canonical isomorphism $\underline{\omega}_E \xrightarrow{\sim} e^*\Omega^1_{E/S}$ [D3, §1].

The first deRham cohomology sheaf of E is defined by

$$\underline{H}^1_{DR}(E) = \mathbb{R}^1\pi_*\Omega^{\cdot}_{E/S}. \tag{1.4}$$

It is locally free of rank 2 on S, and the spectral sequence of hypercohomology gives an exact sequence of sheaves on S [K2, A 1.2]

$$0 \to \underline{\omega}_E \to \underline{H}^1_{DR}(E) \to \underline{\mathrm{Lie}}(E) \to 0. \tag{1.5}$$

The cup-product on deRham cohomology defines an alternating pairing of sheaves:

$$\langle\ ,\ \rangle_{DR}\colon \underline{H}^1_{DR}(E) \times \underline{H}^1_{DR}(E) \to \mathcal{O}_S$$

as $\underline{H}^2_{DR}(E) = R^1\pi_*\Omega^1_{E/S}$, which is isomorphic to $\mathcal{O}_S$ by the trace map. This pairing induces the duality of $\underline{\omega}_E$ and $\underline{\mathrm{Lie}}(E)$. If S is smooth over T, there is a canonical integrable connection (the Gauss-Manin connection [K2, A1.3])

$$\nabla\colon \underline{H}^1_{DR}(E) \to \underline{H}^1_{DR}(E) \otimes \Omega^1_{S/T}.$$

Using this connection, and the cup product, we may define a morphism of sheaves on S:

$$\begin{aligned} i\colon \underline{\omega}_E \otimes \underline{\omega}_E &\to \Omega^1_{S/T} \\ \omega \otimes \nu &\mapsto \langle \omega, \nabla\nu\rangle_{DR}. \end{aligned} \tag{1.6}$$

Let p be a prime, and let S be a scheme over $\mathbb{Z}/p\mathbb{Z}$. If E is an elliptic curve over S, its Hasse invariant $A(E)$ is a section of the invertible sheaf $\underline{\omega}_E^{\otimes p-1}$ [KM, 12.4]. If ω is a non-vanishing section of $\underline{\omega}_E$ over the open set U and C is the (p^{-1}-linear) Cartier operator on differentials, then the restriction of $A(E)$ to U is given by the formula $C(\omega)^p \cdot \omega^{-1}$. The geometric fibers over which $A(E_s) = 0$ are called supersingular, and the fibres where $A(E_s) \neq 0$ are called ordinary.

We illustrate these general notions with a consideration of the Tate curve $E = \mathbb{G}_m/q^{\mathbb{Z}}$, which is an elliptic curve over the ring $\mathbb{Z}((q)) = \mathbb{Z}[[q]][q^{-1}]$ [DR, VII §1]. Here we have an exact sequence of group schemes [KM, 8.8]

$$0 \longrightarrow \mu_N \xrightarrow[Id_N]{} E_N \longrightarrow \mathbb{Z}/N\mathbb{Z} \longrightarrow 0$$

where μ_N is the N-torsion in $\mathbb{G}_m$. The pairing e_N of (1.1) is determined by the formula

$$e_N(\zeta, q^{1/N}) = \zeta \tag{1.8}$$

where ζ is any section of μ_N and $q^{1/N}$ is any N^{th} root of q, i.e., $q^{1/N}$ is any point in E_N mapping to 1 (mod N) in $\mathbb{Z}/N\mathbb{Z}$. Hence the isomorphism f_N of (1.2) induces the identity maps on $\mu_N = \mathrm{Hom}(\mathbb{Z}/N\mathbb{Z}, \mathbb{G}_m)$ and $\mathbb{Z}/N\mathbb{Z} = \mathrm{Hom}(\mu_N, \mathbb{G}_m)$ in (1.7). Concerning the invertible sheaf $\underline{\omega}_E$, we have

PROPOSITION 1.9. a) *The sheaf $\underline{\omega}_E$ has a non-vanishing section dt/t, where t is the parameter on $\mathbb{G}_m$.*

b) *In the map i defined by* (1.6) *the image of the section* $(dt/t)^{\otimes 2}$ *of* $\underline{\omega}_E^{\otimes 2}$ *is the section* (dq/q) *of* $\Omega^1_{\mathbb{Z}((q))/\mathbb{Z}}$.

c) *Over the base* $\mathbb{Z}/p\mathbb{Z}((q))$ *the Hasse invariant of E is given by* $A(E) = (dt/t)^{\otimes p-1}$.

Part a) is proved in [DR, VII, 1.16.2], part b) is proved in [K2, A 1.3.18], and part c) is proved in [KM, 12.4.2]. In the last two references, the differential (dt/t) is called ω_{can}.

A generalized elliptic curve E over the base S [DR, II 1.12] is a scheme of curves $\pi\colon E \to S$ whose geometric fibres are either elliptic curves or Néron polygons, together with a morphism $+\colon E^{reg} \times_S E \to E$ whose restriction to E^{reg} (the union of smooth points in the fibres) makes E^{reg} into a commutative group scheme over S. One insists that the morphism $+$ defines an action of the group scheme E^{reg} on E, and that on the fibres E_s with singular points the translations by E_s^{reg} act by rotations on the graph of irreducible components. One can define the invertible sheaf $\underline{\omega}_E$ on S, for a generalized elliptic curve, as the dual of the sheaf of Lie algebras $\underline{\mathrm{Lie}}(E^{reg})$. The Tate curve $E = \mathbb{G}_m/q^{\mathbb{Z}}$ is a generalized elliptic curve over the base $\mathbb{Z}[[q]]$ [DR, VII, §1]; the line bundle $\underline{\omega}_E$ again has a non-vanishing section dt/t and for $N \geqslant 1$ there is again a canonical homomorphism of group schemes $Id_N\colon \mu_N \to E_N \underset{\text{def}}{=} E_N^{reg}$.

§2. **Modular forms.** This section contains a brief review of the geometric theory of holomorphic modular forms for the group $\Gamma_1(N)$. We refer to the papers of Deligne-Rapoport [DR] and Katz [K2] for the proofs.

Let k and N be integers $\geqslant 1$, and let R be a commutative ring in which N is invertible. A holomorphic modular form f of weight k for $\Gamma_1(N)$, defined over R, is a law which assigns to every pair (E, α)—consisting of a generalized elliptic curve E over an R-algebra A and an embedding of group schemes $\alpha\colon \mu_N \hookrightarrow E_N$ over A whose image meets every irreducible component in each geometric fibre of E—an element $f(E, \alpha) \in \underline{\omega}_E^{\otimes k}$ [DS, §2.1]. This law must be compatible with isomorphisms and extension of scalars. Since all of our modular forms will be holomorphic, we will refer to f simply as a modular form. Let $M_k(R)$ denote the R-module of all modular forms of weight k for $\Gamma_1(N)$.

We reinterpret this definition using the following.

Proposition 2.1. *The functor which assigns to each* $\mathbb{Z}[1/N]$*-scheme S the set of isomorphism classes of pairs* (E, α), *where E is a generalized elliptic curve over S and* $\alpha\colon \mu_N \hookrightarrow E_N$ *an embedding of group schemes whose image meets every irreducible component in each geometric fibre, is represented by an algebraic stack which is proper and smooth over* $\mathbb{Z}[1/N]$. *When* $N > 4$ *this functor is represented by an algebraic curve* $X_1(N)$, *which is proper, smooth, and geometrically connected over* $\mathbb{Z}[1/N]$.

Proof. Let H be the subgroup of $GL_2(\mathbb{Z}/N\mathbb{Z})$ consisting of matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with $c \equiv 0 \pmod N$ and $d \equiv 1 \pmod N$, and let $\mathcal{M}_H[1/N]$ and $\mathcal{M}_H^0[1/N]$ be the algebraic stacks over $\mathbb{Z}[1/N]$ defined in [DR, IV.3]. By definition, $\mathcal{M}_H^0[1/N]$ classifies triples

(E, C, β), where E is an elliptic curve over S, C a subgroup scheme of E_N which is locally isomorphic to $\mathbb{Z}/N\mathbb{Z}$, and β is an isomorphism $\beta\colon \mathbb{Z}/N\mathbb{Z} \overset{\sim}{\rightarrow} E_N/C$. Let $\tilde{\beta}(1)$ be any section of E_N which maps to $\beta(1)$ in E_N/C. By the non-degeneracy of the pairing defined in (1.1), there is a unique isomorphism of group schemes $\alpha\colon \mu_N \overset{\sim}{\rightarrow} C$ such that $e_N(\alpha(\zeta), \tilde{\beta}(1)) = \zeta$ for all sections ζ of μ_N. Hence the data (E, C, β) is equivalent to the data (E, α), and the stack $\mathscr{M}_H^0[1/N]$ classifies these pairs (when E is a genuine elliptic curve). The argument of [DR, IV, 3.5.1] shows that the stack $\mathscr{M}_H[1/N]$ represents this functor for generalized curves.

When $N > 4$ the objects (E, α) classified by $\mathscr{M}_H[1/N]$ have no automorphisms [KM, 2.7.4], so the stack $\mathscr{M}_H[1/N]$ is a projective and smooth scheme over $\mathbb{Z}[1/N]$ [DR, III 2.9]. We denote this scheme by $X_1(N)$, the complex theory [DR, III §5] then shows that $X_1(N)$ is a geometrically connected curve.

We will *henceforth assume that* $N > 4$, so that the stack classifying pairs (E, α) is a scheme, and will treat the cases when $N \leqslant 4$ in §10. Let $\underline{E}$ be the universal family of generalized elliptic curves over $X_1(N)$ (with fixed embedding $\underline{\alpha}\colon \mu_N \hookrightarrow \underline{E}_N$) and let $\underline{\omega} = \underline{\omega}_{\underline{E}}$ be the line bundle on the curve $X_1(N)$ defined at the end of §1 ($\underline{\omega}$ is the dual of the Lie algebra bundle $\underline{\mathrm{Lie}}(\underline{E}^{reg})$).

PROPOSITION 2.2. *The space of modular forms of weight k for $\Gamma_1(N)$ defined over R is equal to $H^0(X_1(N), \underline{\omega}^{\otimes k} \otimes R)$.*

Proof. This is simply a restatement of our definition, using the existence of a universal curve $\underline{E}$ on $X_1(N)$.

We now investigate the line bundle $\underline{\omega}^{\otimes 2}$, using the map defined in (1.6). Let *cusps* denote the divisor on the curve $X_1(N)$ over which the fibres of $\underline{E}$ are Néron polygons, and let $X_1(N)^0$ denote the open curve obtained by removing the divisor *cusps*.

PROPOSITION 2.3. *On the curve $X_1(N)^0$, the map (1.6) of sheaves $i\colon \underline{\omega}^{\otimes 2} \rightarrow \Omega^1_{X_1(N)^0/\mathbb{Z}[1/N]}$ is an isomorphism. This extends to an isomorphism of sheaves*

$$\underline{\omega}^{\otimes 2} \overset{i}{\underset{\sim}{\rightarrow}} \Omega^1_{X_1(N)}(cusps) \quad on\ X_1(N). \tag{2.4}$$

Proof. The two statements are proved in [K2, A 1.3.17] for any universal family $\underline{E}$ of elliptic curves. The first is a consequence of the Kodaria-Spencer theory of deformations, and the second follows from a calculation on Tate curves (cf. Proposition 1.9, b).

Let g be the genus of the (geometrically connected) curve $X_1(N)$. Formula (2.4) shows that $\deg(\underline{\omega}^{\otimes k}) \geqslant 2g - 1$ for $k \geqslant 2$, so $H^1(X_1(N), \underline{\omega}^{\otimes k}) = 0$ for all $k \geqslant 2$ by Serre duality. As a corollary of this fact, one obtains

PROPOSITION 2.5. (*cf.* [K2, 1.7.1])
For $k \geqslant 2$, the natural map $H^0(X_1(N), \underline{\omega}^{\otimes k}) \otimes R \rightarrow H^0(X_1(N), \underline{\omega}^{\otimes k} \otimes R)$ is an isomorphism.

Note. The map of (2.5) need *not* be an isomorphism when $k = 1$ and $R = \mathbb{Z}/p\mathbb{Z}$ [S9].

We end by defining the Fourier expansion of a modular form f defined over R at the cusp ∞ of $X_1(N)$. Recall that the Tate curve $E = \mathbb{G}_m/q^{\mathbb{Z}}$ is a generalized elliptic curve over $\mathbb{Z}[[q]]$, which has a canonical differential dt/t as well as a natural embedding $Id_N\colon \mu_N \hookrightarrow E_N$ over $\mathbb{Z}[1/N][[q]]$. We define the Fourier expansion $f(q) = \Sigma_{n \geqslant 0} a_n q^n$ of f in the ring $R[[q]]$ by the formula

$$f(\mathbb{G}_m/q^{\mathbb{Z}}, Id_N) = f(q)\cdot(dt/t)^{\otimes k}.$$

Since $X_1(N)$ represents the functor of pairs (E, α), there is a unique morphism $Spec\, \mathbb{Z}[1/N][[q]] \to X_1(N)$ over $\mathbb{Z}[1/N]$ such that $(\mathbb{G}_n/q^{\mathbb{Z}}, Id_N)$ arises from pull-back of the universal pair $(\underline{E}, \underline{\alpha})$. The image of the prime ideal where $q = 0$ defines the section ∞ of $X_1(N)$, and q is a uniformizing parameter in the neighborhood of this cusp. Hence the Fourier expansion $f(q)$ describes the holomorphic section f of $\underline{\omega}^{\otimes k}$ in the neighborhood of ∞. Since $X_1(N)$ is geometrically connected, we find:

PROPOSITION 2.7. (*cf.* [K2, 1.6.1, 1.6.2])

a) *The map* $H^0(X_1(N), \underline{\omega}^{\otimes k} \otimes R) \to R[[q]]$ *taking* f *to* $f(q)$ *is an injection of* R*-modules.*

b) *If* R_0 *is a sub* $\mathbb{Z}[1/N]$*-algebra of* R, *the modular form* f *is defined over* R_0 *if and only if* $f(q) \in R_0[[q]]$.

Using the isomorphism (2.4), we may identify a modular form f of weight 2 with a meromorphic differential ω_f on $X_1(N)$, which is regular outside *cusps* and has poles of order $\leqslant 1$ along each cuspidal section.

PROPOSITION 2.8. *The expansion of* ω_f *in a neighborhood of the cusp* ∞ *is given by* $f(q)dq/q$.

Proof. This follows from part b) of (1.9), which shows that the local section $(dt/t)^{\otimes 2}$ of $\underline{\omega}^{\otimes 2}$ is mapped to the local differential dq/q of $\mathbb{Z}((q))$.

§3. Hecke operators. In this section we define certain endomorphisms—the Hecke operators T_l and U_l and the automorphisms $\langle d\rangle$—of the space of modular forms of weight k for $\Gamma_1(N)$ over R. We also discuss their action as correspondences of the curve $X_1(N)$ over $\mathbb{Z}[1/N]$.

Let d be a class in $(\mathbb{Z}/N\mathbb{Z})^{\times}$, and define the automorphism (d) of $X_1(N)$ over $\mathbb{Z}[1/N]$ by:

$$\langle d\rangle(E, \alpha) = (E, d\alpha). \tag{3.1}$$

Here, as usual, E is a generalized elliptic curve over a scheme S where N is invertible and $\alpha\colon \mu_N \hookrightarrow E_N$ is an embedding of group schemes whose image meets every irreducible component in each geometric fibre of E. The embedding $d\alpha$ maps the section ζ of μ_N to $d\cdot\alpha(\zeta)$ in E_N. The automorphism $\langle d\rangle$ induces an R-linear automorphism $f \mapsto f|\langle d\rangle$ of the space of modular forms of weight k for $\Gamma_1(N)$ over

R; we have

$$f|\langle d\rangle\,(E,\alpha) = f(E, d\alpha) \quad \text{in } \underline{\omega}_E^{\otimes k}. \tag{3.2}$$

If $\varepsilon\colon (\mathbb{Z}/N\mathbb{Z})^\times \to R^\times$ is a group homomorphism with $\varepsilon(-1) = (-1)^k$ and f is a modular form of weight k over R, we say f has type (k, ε) if $f|\langle d\rangle = \varepsilon(d)\cdot f$ for all $d \in (\mathbb{Z}/N\mathbb{Z})^\times$.

Let l be a prime with $l \nmid N$, and let f be a modular form of weight k for $\Gamma_1(N)$ over $\mathbb{Z}[1/N]$. We first define $f|T_l$ as a modular form of weight k over $\bar{\mathbb{Q}}$, the algebraic closure of $\mathbb{Q}$. If E is an elliptic curve over $\bar{\mathbb{Q}}$ and $\alpha\colon \mu_N \hookrightarrow E_N$ is an embedding, define

$$f|T_l\,(E,\alpha) = \frac{1}{l}\sum_{\varphi} \varphi^*(f(\varphi E, \varphi\alpha)), \tag{3.3}$$

where the sum is taken over the $(l + 1)$ isogenies $\varphi\colon E \to \varphi E$ of degree l with source E. A calculation on the Tate curve [cf. K2, 1.11] shows that the law $f|T_l$ extends uniquely to generalized elliptic curves, so defines a modular form of weight k over $\bar{\mathbb{Q}}$. If f and $f|\langle l\rangle$ have Fourier expansions

$$\begin{cases} f(q) = \sum a_n q^n \\ f|\langle l\rangle(q) = \sum b_n q^n \end{cases}$$

at ∞, then $f|T_l$ has the Fourier expansion [K2, 1.11.2]:

$$f|T_l(q) = \sum a_{nl} q^n + l^{k-1} \sum b_n q^{nl}. \tag{3.5}$$

Since f and $f|\langle l\rangle$ are defined over $\mathbb{Z}[1/N]$, the coefficients of $f|T_l$ belong to the subring $\mathbb{Z}[1/N]$ of $\bar{\mathbb{Q}}$. Hence $f|T_l$ is a modular form of weight k over $\mathbb{Z}[1/N]$, by Proposition 2.7 b). If l is a prime dividing N, we define $f|U_l$ as a modular form of weight k over $\bar{\mathbb{Q}}$ by the formula

$$f|U_l\,(E,\alpha) = \frac{1}{l}\sum_{\varphi} \varphi^*(f(\varphi E, \varphi\alpha)) \tag{3.6}$$

where E is an elliptic curve and the sum is now taken over the l isogenies $\varphi\colon E \to \varphi E$ of degree l, whose kernel has trivial intersection with the image of α in E_N. Again this law extends uniquely to generalized curves, and has Fourier expansion

$$f|U_l(q) = \sum a_{nl} q^n. \tag{3.7}$$

Proposition 2.7 b) again shows that $f|U_l$ is defined over $\mathbb{Z}[1/N]$.

Let R be a $\mathbb{Z}[1/N]$-algebra. For $k \geqslant 2$ we have an isomorphism:

$$H^0(X_1(N), \underline{\omega}^{\otimes k}) \otimes R \xrightarrow{\sim} H^0(X_1(N), \underline{\omega}^{\otimes k} \otimes R)$$

by Proposition 2.5. Hence the operators T_l and U_l define (by extension of scalars) endomorphisms of the R-module of modular forms of weight $k \geqslant 2$ for $\Gamma_1(N)$ over R (cf. [K2, 1.11.4]). When $k = 1$ this argument does not apply. Nevertheless, the proof sketched above for $R = \mathbb{Z}[1/N]$ works for weight $k = 1$ over any subring R of $\bar{\mathbb{Q}}$. We will treat the operators T_l and U_l on modular forms of weight 1 over $R = \mathbb{Z}/p\mathbb{Z}$ in the next section.

The endomorphisms T_l (for $l \nmid N$), U_l (for $l \mid N$), and $\langle d \rangle$ (for $d \in (\mathbb{Z}/N\mathbb{Z})^\times$) of the space of modular forms of weight k over R all commute with each other. If f has type (k, ε) then formula (3.5) simplifies to:

$$f \mid T_l(q) = \sum a_{nl} q^n + \varepsilon(l) l^{k-1} \sum a_n q^{nl}. \tag{3.8}$$

We say f is a normalized eigenform of type (k, ε) if it is a simultaneous eigenvector for the operators T_l and U_l, for all primes l, and satisfies $a_1 = 1$. We then have the following formula for the higher Fourier coefficients of a normalized eigenform:

$$f \mid T_l = a_l \cdot f \quad \text{all } l \nmid N$$

$$f \mid U_l = a_l \cdot f \quad \text{all } l \mid N$$

$$\sum_{n \geqslant 1} a_n n^{-s} = \prod_{l \mid N} (1 - a_l l^{-s})^{-1} \prod_{l \nmid N} (1 - a_l l^{-s} + \varepsilon(l) l^{k-1-2s})^{-1}. \tag{3.9}$$

We say f is a cusp form if, as a section of $\underline{\omega}^{\otimes k}$ over $X_1(N)$, it vanishes along the divisor *cusps*. In particular, this implies that f vanishes at ∞ and $a_0 = 0$. By (3.9) the entire Fourier expansion of a normalized cuspidal eigenform f is determined by its character ε and its set of eigenvalues (a_l). We say f is a "new form" for $\Gamma_1(N)$ over R if f is a normalized cuspidal eigenform of weight k whose set of eigenvalues $\{a_l \colon l \nmid N\}$ does *not* occur for an eigenform of weight k for $\Gamma_1(M)$ over R, for any proper divisor M of N. When $R = \mathbb{C}$, this is equivalent to the definition of [AL].

We now define the Hecke correspondences T_l and U_l of the curve $X_1(N)$ over $\mathbb{Z}[1/N]$. First assume that $l \nmid N$, and let $X_1(N; l)$ be the fine moduli scheme over $\mathbb{Z}[1/N]$ which represents the functor of triples (E, α, C), where E is a generalized elliptic curve, $\alpha \colon \mu_N \hookrightarrow E_N$ an embedding, and C a locally free subgroup scheme of rank l in E_l. One insists further that the finite group scheme Image $\alpha \times C$ meets every irreducible component in each geometric fibre of E [DR, V 1.6]. If E is a genuine elliptic curve, let $E' = E/C$ and let $\varphi \colon E \to E'$ be the associated l-isogeny. We define morphisms π_1 and π_2 of schemes over $\mathbb{Z}[1/N]$:

$$\begin{cases} \pi_1 \colon X_1(N; l) \to X_1(N) \\ \quad (E, \alpha, C) \mapsto (E, \alpha) \\ \pi_2 \colon X_1(N; l) \to X_1(N) \\ \quad (E, \alpha, C) \mapsto (E', \alpha' = \varphi\alpha). \end{cases} \tag{3.10}$$

Strictly speaking, these maps are only defined on the non-cuspidal points, but they extend uniquely to $X_1(N; l)$. They are both finite coverings of degree $l + 1$, and are étale coverings of $X_1(N)^0$ over $\mathbb{Z}[1/Nl]$. We define the correspondence T_l (cf. [Sh2, 7.3]) as the image of $X_1(N; l)$ in $X_1(N) \times_{\mathbb{Z}[1/N]} X_1(N)$ under the map $\pi_1 \times \pi_2$.

The correspondence T_l acts on divisors d of $X_1(N)$ by the formula: $T_l(d) = pr_2(T_l \cdot (d \times X_1(N))$. If $P = (E, \alpha)$ is a non-cuspidal point, we find

$$(3.11) \qquad T_l(E, \alpha) = \sum_{\varphi} (\varphi E, \varphi\alpha),$$

where the sum is taken over the l-isogenies with source E. The action on divisors preserves the subgroup of principal divisors, so induces an endomorphism (also denoted by T_l) of the Jacobian $J_1(N)$ over $\mathbb{Z}[1/N]$. If we extend the formula (3.11) to divisors of degree zero modular linear equivalence, this is the action on $J_1(N)$ induced by Albanese functorality. Since $l \nmid N$ we may consider the action on $X_1(N)$ or $J_1(N)$ over $\mathbb{Z}/l\mathbb{Z}$.

PROPOSITION 3.12. (*cf.* [E], [Sh1], [Sh2, Thm. 7.9] *Let Fr_l be the Frobenius correspondence of $X_1(N)$ over $\mathbb{Z}/l\mathbb{Z}$ and let $Ver_l = {}^tFr_l$ be its transpose. Then $T_l \equiv Ver_l + \langle l \rangle Fr_l$ in the ring of correspondences of the curve $X_1(N)$ over $\mathbb{Z}/l\mathbb{Z}$, and in the endomorphism ring of the Jacobian $J_1(N)$ over $\mathbb{Z}/l\mathbb{Z}$.*

Proof. If (E, α) is an ordinary point on $X_1(N)^0$ in characteristic l, one checks that $T_l(E, \alpha) = Ver_l(E, \alpha) + Fr_l(E, l\alpha)$. Since the ordinary points are dense, this verifies the claim.

We warn the reader that many authors [e.g., MW] prefer to work with the curve $X_1(N)'$, which classifies pairs (E, a) where $a: \mathbb{Z}/N\mathbb{Z} \hookrightarrow E_N$ is an embedding. On the curve $X_1(N)'$, the congruence in Proposition 3.12 becomes $T_l' \equiv Fr_l' + \langle l \rangle' Ver_l'$ (mod l).

For $l | N$ we may define U_l as a correspondence on $X_1(N)$ over $\mathbb{Z}[1/N]$ by considering the fine moduli scheme $X_1(N; l)$ over $\mathbb{Z}[1/N]$ which classifies triples (E, α, C) as before, with the extra condition that $C \cap Im\, \alpha = 0$. (This is the stack associated to the subgroup H of $GL_2(\mathbb{Z}/N\mathbb{Z})$ of matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with $c \equiv 0 \pmod{N}$, $d \equiv 1 \pmod{N}$, and $b \equiv 0 \pmod{l}$.) Again there are two natural coverings $\pi_1, \pi_2: X_1(N; l) \to X_1(N)$ and we define U_l as the image of $\pi_1 \times \pi_2$ in the product $X_1(N) \times_{\mathbb{Z}[1/N]} X_1(N)$. The action of U_l on a point (E, α) of $X_1(N)^0$ is given by

$$(3.13) \qquad U_l(E, \alpha) = \sum_{\varphi} (\varphi E, \varphi\alpha)$$

where the sum is taken over the l-isogenies φ with source E such that $\varphi\alpha: \mu_N \to \varphi E$ is an embedding.

The correspondences T_l and U_l act on the holomorphic differentials ω on $X_1(N)$ by the formula $\omega | T_l = (\pi_1)_* \circ \pi_2^* \omega$, where π_2^* is the pull-back and $(\pi_1)_*$ is the trace map associated to a finite covering of proper curves. If we identify $H^0(X_1(N), \Omega^1_{X_1(N)})$ with the invariant differentials on $J_1(N)$, the action of T_l and U_l on holomorphic differentials is the one induced from their action as endomorphisms of the Albanese variety (cf. the discussion, without proofs, in [MW Ch. 2, §5.4, §5.8]). At the point (E, α) of $X_1(N)^0$ we find

$$\omega | T_l(E, \alpha) = \sum_{\varphi} \omega(\varphi E, \varphi\alpha) \tag{3.14}$$

where the sum is over the l-isogenies of source E and the fibres of $\Omega^1_{X_1(N)\mathbb{Z}[1/Nl]}$ at E and φE are identified using the maps π_1 and π_2.

Recall the isomorphism $i: \underline{\omega}^{\otimes 2} \xrightarrow{\sim} \Omega^1_{X_1(N)}$ (*cusps*) of invertible sheaves on $X_1(N)$ which was defined in (2.4). This induces an isomorphism of global sections, and hence an isomorphism from the space $H^0(X_1(N), \underline{\omega}^{\otimes 2})^0$ of cusp forms of weight 2 over $\mathbb{Z}[1/N]$ to the space $H^0(X_1(N), \Omega^1_{X_1(N)})$ of holomorphic differentials over $\mathbb{Z}[1/N]$. We have defined endomorphisms T_l, U_l and $\langle d \rangle$ of both $\mathbb{Z}[1/N]$-modules; it remains to check the following.

PROPOSITION 3.15. (*cf.* [Sh2, 7.2.6])
The map $i: H^0(X_1(N), \underline{\omega}^{\otimes 2})^0 \xrightarrow{\sim} H^0(X_1(N), \Omega^1_{X_1(N)})$ *is an isomorphism of Hecke modules: it commutes with the action of* T_l, U_l, *and* $\langle d \rangle$.

Proof. Let $v \otimes v'$ be a local section of $\underline{\omega}^{\otimes 2}$ on the set $U \subset X_1(N)^0$, where v and v' are relative differentials on the universal curve over U. Then $i(v \otimes v') = \langle v, \nabla v' \rangle_{DR}$ as a regular differential on U over $\mathbb{Z}[1/N]$. Assume that the point (E, α) and its translates $(\varphi E, \varphi\alpha)$ by T_l are contained in U. By the definition (3.3) of the endomorphism T_l acting on forms of weight 2:

$$\begin{aligned}(v \otimes v') | T_l(E, \alpha) &= \frac{1}{l} \sum_{\varphi} \varphi^*(v_{\varphi E} \otimes v'_{\varphi E}) \\ &= \frac{1}{l} \sum_{\varphi} \varphi^* v_{\varphi E} \otimes \varphi^* v'_{\varphi E}.\end{aligned}$$

Hence the image of $(v \otimes v') | T_l$ under i is equal to the differential

$$\omega(E, \alpha) = \frac{1}{l} \sum_{\varphi} \langle \varphi^* v_{\varphi E}, \nabla \varphi^* v'_{\varphi E} \rangle_{DR}, \tag{3.16}$$

where, as usual, the fibres of $\Omega^1_{X_1(N)}$ at (E, α) and $(\varphi E, \varphi\alpha)$ have been identified.

But if $\varphi: E \to F$ is an isogeny and v and v' are invariant differentials on F, we have

$$\nabla_E \varphi^* v' = \varphi^* \nabla_F v' \quad \text{in } \underline{H}^1_{DR}(E)$$

$$\langle \varphi^* v, \varphi^* \nabla v' \rangle^E_{DR} = \deg \varphi \cdot \langle v, \nabla v' \rangle^F_{DR}.$$

The first follows from the naturality of the Gauss-Manin connection and the second from the fact that the adjoint of the isogeny φ with respect to the pairing $\langle\ ,\ \rangle_{DR}$ is the isogeny ${}^t\varphi$, which satisfies ${}^t\varphi \circ \varphi = \deg \varphi$. Since all the isogenies φ in the sum of (3.16) have $\deg \varphi = l$, we find that

$$\omega(E, \alpha) = \sum_{\varphi} \langle v_{\varphi E}, \nabla v'_{\varphi E} \rangle_{DR}.$$

This differential is equal to $i(v \otimes v')|T_l$, under the action (3.14) of T_l on the space of holomorphic differentials on $X_1(N)$. Hence $i(v \otimes v')|T_l = i(v \otimes v'|T_l)$. A similar proof works for U_l and $\langle d \rangle$.

Note. R. Coleman has observed that Proposition 3.15 also follows from the formulas: $i(\pi_1^* f) = \pi_1^*(if)$, $i(\pi_2^* f) = l \cdot \pi_2^*(if)$ where π_1 and π_2 are the maps defined in (3.10).

COROLLARY 3.17. *If $\omega = \Sigma_{n \geqslant 1} a_n q^n dq/q$ is the formal expansion of the holomorphic differential ω in a neighborhood of the cusp ∞ and $\omega|\langle l \rangle = \Sigma_{n \geqslant 1} b_n q^n dq/q$, then*

$$\omega|T_l = \left(\sum_{n \geqslant 1} a_{nl} q^n + l \sum b_n q^{nl} \right) dq/q$$

$$\omega|U_l = \sum_{n \geqslant 1} a_{nl} q^n dq/q.$$

Proof. This follows from a combination of Proposition 3.15, Proposition 2.8 (which relates the local expansion of $\omega = \omega_f$ with the Fourier expansion of the cusp form f of weight 2) and formulas (3.5) and (3.7) (which give the action of T_l and U_l on q-expansions).

We remark that the first formula in (3.17) can be used to give a different proof of the Eichler-Shimura congruence in Proposition 3.12.

§4. Modular forms (mod p). We henceforth fix a prime p which does not divide N (and recall that $N \geqslant 4$). For $k \geqslant 1$ we let M_k denote the vector space of modular forms of weight k for $\Gamma_1(N)$ over $\mathbb{Z}/p\mathbb{Z}$, and let M_k^0 denote the space of cusp forms. If R is a field of characteristic p, then the space of modular forms of weight k over R is $M_k \otimes R$. We let σ denote the Frobenius endomorphism of R over $\mathbb{Z}/p\mathbb{Z}$: $\sigma(x) = x^p$. If f in $M_k \otimes R$ has Fourier expansion $\Sigma\, a_n q^n$, then $f^\sigma(q) = \Sigma\, a_n^\sigma q^n = \Sigma\, a_n^p q^n$.

In the previous section, we defined linear endomorphisms T_l, U_l, and $\langle d \rangle$ of M_k and M_k^0, provided that $k \geqslant 2$. In this case it is customary to denote the operator T_p by U_p—this makes little difference as the formulae (3.5) and (3.7) agree on q-expansions. We now consider the case when $k = 1$. For $l \neq p$, formula (3.3) defines an endomorphism of the space of forms over an algebraic closure of $\mathbb{Z}/p\mathbb{Z}$, and the q-expansion (3.5) shows that T_l gives an endomorphism of M_k stabilizing M_k^0. A similar argument, using (3.6) and the q-expansion (3.7) gives the existence of U_l, for $l|N$. The only difficulty remaining is to define the operator T_p.

PROPOSITION 4.1. *There is a unique endomorphism T_p: $M_1 \to M_1$ which has the following effect on q-expansions: if $f \in M_1$ and $f(q) = \Sigma\, a_n q^n$, $f|\langle p\rangle(q) = \Sigma\, b_n q^n$, then*

$$f|T_p(q) = \sum a_{np} q^n + \sum b_n q^{np}.$$

Proof. By part a) of Proposition 2.7, the form $f|T_p$ (if it exists) of weight 1 is completely determined by its Fourier expansion at ∞. Hence we are reduced to proving that $\Sigma\, a_{np} q^n + \Sigma\, b_n q^{np}$ is the Fourier expansion of a holomorphic form g of weight 1 over $\mathbb{Z}/p\mathbb{Z}$.

The curve $X_1(N)^0 = X_1(N) - cusps$ is affine, so $H^1(X_1(N)^0, \underline{\omega}) = 0$. Hence the natural map $H^0(X_1(N)^0, \underline{\omega}) \otimes \mathbb{Z}/p\mathbb{Z} \to H^0(X_1(N)^0, \underline{\omega} \otimes \mathbb{Z}/p\mathbb{Z})$ is an isomorphism, and we may lift f to a section F of $\underline{\omega}$ on $X_1(N)^0$ over $\mathbb{Z}[1/N]$. F is then a meromorphic modular form of weight 1 for $\Gamma_1(N)$ over $\mathbb{Z}[1/N]$ with singularities along the divisor *cusps*. Similarly $F|\langle p\rangle$ is a meromorphic form, which reduces (mod p) to $f|\langle p\rangle$.

The definition of Fourier expansions at ∞, using the Tate curve $\mathbb{G}_m/q^{\mathbb{Z}}$ over $\mathbb{Z}[1/N]((q))$ as in (2.6), extends to meromorphic modular forms. We have $F(q) = \Sigma\, A_n q^n$ and $F|\langle p\rangle(q) = \Sigma\, B_n q^n$ with $A_n = B_n = 0$ for $n \ll 0$. Since the map taking a meromorphic form to its Fourier expansion commutes with reduction (mod p), we have the congruences: $A_n \equiv a_n$ (mod p) for all $n \geqslant 0$, $A_n \equiv B_n \equiv 0$ (mod p) for all $n < 0$.

The definition of T_p in (3.3), and the Fourier expansion (3.5) of $F|T_p$ also works for meromorphic forms F. Hence $F|T_p(q) = \Sigma\, A_{np} q^n + \Sigma\, B_n q^{np}$ is the Fourier expansion of a meromorphic form of weight 1 over $\mathbb{Z}[1/N]$. Let g be the reduction of $f|T_p$, which is *a priori* a meromorphic form of weight 1 for $\Gamma_1(N)$ over $\mathbb{Z}/p\mathbb{Z}$. Since the negative Fourier coefficients of $F|T_p$ are all $\equiv 0$ (mod p), the Fourier expansion of g at ∞ is equal to $\Sigma_{n\geqslant 0}\, a_{np} q^n + \Sigma_{n\geqslant 0}\, b_n q^{np}$. Hence g is regular at ∞. A similar argument (using the p-divisibility of the negative coefficients of F at all cusps of $X_1(N)$, and the formula for the expansion of $F|T_p$ at other cusps (cf. [K2, 1.11.1])) shows that g is regular at all cusps of $X_1(N)$ over $\mathbb{Z}/p\mathbb{Z}$, so is a holomorphic form of weight 1.

Beyond the Hecke operators on M_k, there are additional linear maps between the spaces which exist only in characteristic p. First there is the map:

(4.2)
$$V_p\colon M_k \to M_{pk}$$
$$f \mapsto f^p.$$

The linear extension of V_p to $M_k \otimes R$ satisfies

(4.3) $$(V_p f)^\sigma = V_p(f^\sigma) = f^p,$$

and the effect of V_p on Fourier expansions is given by

(4.4) $$f|V_p(q) = \sum a_n q^{np}.$$

Next, there is the derivation:

$$\theta\colon M_k \to M_{k+(p+1)} \tag{4.5}$$

which is characterized by its action on Fourier expansions:

$$\theta f(q) = \sum n a_n q^n. \tag{4.6}$$

The existence of θ is proved as in Katz [K3, where θ is denoted $A\theta$]; our hypothesis that $N \geqslant 4$ insures the existence of a universal curve $\underline{E}$ over $X_1(N)$. The map θ is injective if k is prime to p, and the kernel of θ on M_{pk} is equal to the image of V_p [K3, §II].

The Hasse invariant $A(E) \in H^0(X_1(N), \underline{\omega}^{\otimes p-1})$ defined in §1 gives a canonical holomorphic modular form A of weight $p-1$ for $\Gamma_1(N)$ over $\mathbb{Z}/p\mathbb{Z}$. By Proposition 1.9 part c) the Fourier expansion of A at the cusp ∞ is given by $A(q) = 1$. The section A vanishes to order 1 at each super-singular point of $X_1(N)$ [KM, 12.4.3]. Multiplication by A gives an injective linear map $f \mapsto Af$ from M_k to $M_{k+(p-1)}$, whose image consists of the sections of $\omega^{\otimes k+(p-1)}$ which vanish along the divisor of supersingular points on $X_1(N)$.

The endomorphisms T_l, U_l, T_p ($= U_p$ for $k \geqslant 2$), and $\langle d \rangle$ of M_k all commute. They also commute with multiplication by $A\colon M_k \to M_{k+(p-1)}$, except in the case when $k = 1$, where:

$$A(f | T_p) = (Af) | U_p + (f | \langle p \rangle) | V_p. \tag{4.7}$$

(This identity suggests a different proof of Proposition 4.1, which does not involve lifting to meromorphic forms in characteristic zero. Namely, for $f \in M_1$ the right hand side of (4.7) defines an element of M_p with the correct Fourier expansion at ∞. A detailed analysis of this section of $\underline{\omega}^{\otimes p}$ shows that it vanishes at each super-singular point, so has the form Ag for $g \in M_1$, and we define $g = f | T_p$.) The operators T_l and U_l commute with $V_p\colon M_k \to M_{pk}$, and $(f | V_p) | U_p = A^k f$. Their commutation relations with θ are:

$$\begin{aligned} (\theta f) | T_l &= l \cdot \theta(f | T_l) \\ (\theta f) | U_l &= l \cdot \theta(f | U_l) \\ (\theta f) | U_p &= \theta(f | V_p) = 0. \end{aligned} \tag{4.8}$$

All this can be checked on Fourier expansions, using part a) of Proposition 2.7.

Let $M = \bigoplus_{k \geqslant 0} M_k$, where $M_0 = \mathbb{Z}/p\mathbb{Z}$. Then M is a graded $\mathbb{Z}/p\mathbb{Z}$-algebra of Krull dimension 2.

Proposition 4.9. (*cf.* [Sw], [S4, Thm. 1]) *The kernel of the ring homomorphism $M \to \mathbb{Z}/p\mathbb{Z}[[q]]$, taking $f \in M_k$ to its Fourier expansion $f(q)$ at ∞, is equal to the principal ideal $(A-1)M$.*

Proof. The kernel clearly contains $(A-1)M$. This is a prime ideal $\mathfrak{p}$ of M, as A has simple zeroes at the supersingular points. Since the image $\tilde{M}$ is not finite, it has dimension 1 and the kernel cannot properly contain $\mathfrak{p}$ for dimension reasons.

By Proposition 4.9 the image $\tilde{M} = M/(A-1)M$ in $\mathbb{Z}/p\mathbb{Z}[[q]]$ is graded by $\mathbb{Z}/(p-1)\mathbb{Z}$: we write $\tilde{M} = \bigoplus_\alpha \tilde{M}_\alpha$ with α in $\mathbb{Z}/(p-1)\mathbb{Z}$. We say the series $\Sigma\, a_n q^n$ in $\tilde{M}_\alpha$ has filtration $k \geqslant 0$ if it is the image of an element $f \in M_k$ which does not vanish at at least one supersingular point of $X_1(N)$. Equivalently, $\Sigma\, a_n q^n$ has filtration k if it is the Fourier expansion $f(q)$ of $f \in M_k$ which is not divisible by A in M. One has $k \equiv \alpha \pmod{p-1}$, by the definition of the grading on $\tilde{M}$.

Proposition 4.10. (*cf.* [Sw, Lemma 5], [J, §7])

a) *If $f(q) = \Sigma\, a_n q^n$ has filtration k and $(k, p) = 1$, then $\theta f(q) = \Sigma\, n a_n q^n$ has filtration $k + p + 1$.*

b) *Assume $f(q) = \Sigma\, a_n q^n$ has filtration k, with $2 \leqslant k \leqslant p$. Define $k' = p + 1 - k$, so $1 \leqslant k' \leqslant p - 1$. Then $\theta^{k'} f(q)$ has filtration $\leqslant p + 1 + k'$, with equality holding if and only if $f \,|\, U_p \neq 0$.*

Proof. a) is proved for $N = 1$ in [Sw], and in [K3] a proof is given for modular forms on the curve $X(N)$ for $N \geqslant 3$. The latter proof generalizes to $X_1(N)$ for $N \geqslant 4$, using the universal elliptic curve over $X_1(N)^0$ and the Gauss-Manin connection on its deRham cohomology.

To prove b), we first note that $\theta^{k'-1} f(q)$ has filtration pk'; this follows from successive applications of a). Indeed, for $i < k' - 1$, $\theta^i f(q)$ has filtration $= k + i(p+1)$ prime to p, and $k + (k'-1)(p+1) = pk'$. To determine the filtration of $\theta^{k'} f(q)$, we use the formula [K3, pg. 8]:

$$\theta f = A\partial f + kBf \tag{4.11}$$

for $f \in M_k$. Here $\partial f \in M_{k+2}$, $A \in M_{p-1}$, and $B \in M_{p+1}$ is a canonical form (the negative of the form denoted B in [K3]) which is non-zero at all supersingular points. We have the identities: $\partial A = B$ and $\partial B = -QA$ where Q is the normalized Eisenstein series of weight 4 and level 1 [S4, Thm. 5]. If we recursively define $f^{(0)} = f$, $f^{(1)} = \partial f$, $f^{(\nu)} = \partial f^{(\nu-1)} - (k + \nu - 2)(\nu - 1)Q \cdot f^{(\nu-2)}$, then $f^{(\nu)}(q)$ has filtration $\leqslant k + 2\nu$. By induction on n, one proves the formula [Sw, pg. 31]:

$$\theta^n f = \sum_{\nu=0}^{n} \frac{n!}{\nu!} \frac{(k+n-1)!}{(n-\nu)!(k+\nu-1)!} A^\nu B^{n-\nu} f^{(\nu)}$$

for all $n \geqslant 0$, starting with (4.11)—which is the case $n = 1$. When $k \leqslant p$ and $n = k' = p + 1 - k$, we have $(k + n - 1) = p$ and the only nonzero term in the above sum occurs when $v = n$: $\theta^{k'}f = A^{k'}f^{(k')}$. Hence the filtration of $\theta^{k'}f(q)$ is equal to the filtration of $f^{(k')}(q)$, which is $\leqslant k + 2(k') = k' + p + 1$.

Using a) we now observe that the filtration of $\theta^{k'}f(q)$ is equal to $k' + p + 1$ if and only if the filtration of $\theta^{p-1}f(q)$ is equal to pk. But

$$\theta^{p-1}f(q) = \sum_{(n,p)=1} a_n q^n = f(q) - f\,|\,U_p|\,V_p(q).$$

If $f\,|\,U_p \neq 0$, it has filtration k, and $f\,|\,U_p|\,V_p(q)$ has filtration pk (as V_p is essentially the p^{th} power map). If $f\,|\,U_p = 0$, the right hand side has filtration k, so $\theta^{k'}f(q)$ has filtration $\leqslant k' + 2$.

If $f(q) = \Sigma\, a_n q^n$ is an element of $\mathbb{Z}/p\mathbb{Z}[[q]]$ we define $f(q)|\,U_p = \Sigma\, a_{np} q^n$. The next result gives a useful criterion for an element in $\tilde{M}_\alpha$ to have low filtration.

PROPOSITION 4.12. (*cf.* [S5, Thm. 6]) *If $f(q) \in \tilde{M}_\alpha$ satisfies $f(q)|\,U_p = \lambda \cdot f(q)$ with $\lambda \neq 0$, then $f(q)$ has filtration k with $2 \leqslant k \leqslant p + 1$.*

Proof. Let k be the filtration of $f(q)$, and let $f \in M_k$ be a form with this Fourier expansion. Then

$$(f\,|\,U_p)|\,V_p(q) = \sum a_{np} q^{np} = f(q) - \theta^{p-1}f(q).$$

The left hand side is the Fourier expansion of $\lambda \cdot f^p$, which lies in M_{pk} and, by hypothesis, has filtration pk. The right hand side has filtration $\leqslant k + (p-1)(p+1)$. Hence $pk \leqslant k + (p-1)(p+1)$, which implies that $k \leqslant p + 1$. If f has weight $k = 1$, so does $g = f\,|\langle p\rangle$ and $f\,|\,T_p = f\,|\,U_p + g\,|\,V_p$. Since $g\,|\,V_p$ has filtration p, so does $f\,|\,U_p$ and we cannot have $f\,|\,U_p = \lambda f$. Hence $k \geqslant 2$.

The results in Propositions 4.9, 4.10, and 4.12 also hold for modular forms $f \in M_k \otimes R$ and for Fourier expansions $f(q) \in \tilde{M}_\alpha \otimes R \subseteq R[[q]]$, where R is any field of characteristic p. The proofs are essentially the same.

§5. Igusa curves. Let $X_1(N)^h$ be the affine curve over $\mathbb{Z}/p\mathbb{Z}$ obtained by removing the supersingular points (the support of the divisor of the section A of $\underline{\omega}^{\otimes p-1}$) [DR, V, pg. 101]. We define the affine Igusa curve $I_1(N)^h$ as the fine moduli space of triples (E, α, β), where E is a generalized elliptic curve over a scheme of characteristic p, α: $\mu_N \hookrightarrow E_N$ is an embedding whose image meets every irreducible component in each geometric fibre of E, and β: $\mu_p \hookrightarrow E_p$ is an embedding of group schemes.

The group $(\mathbb{Z}/p\mathbb{Z})^\times$ acts freely on $I_1(N)^h$. If d is a non-zero class (mod p) the automorphism $\langle d\rangle_p$ is defined by the formula:

$$\langle d\rangle_p(E, \alpha, \beta) = (E, \alpha, d\beta). \tag{5.1}$$

The quotient of $I_1(N)^h$ by this action is the affine curve $X_1(N)^h$.

Let $\underline{E}$ be the universal curve over $I_1(N)^h$, and let $\underline{\omega}^h$ be the dual of the invertible sheaf $\mathrm{Lie}(\underline{E}^{reg})$. Then $\underline{\omega}^h = \pi^*(\underline{\omega})$, where π: $I_1(N)^h \to X_1(N)^h$ is the covering map.

There is a unique morphism Spec $\mathbb{Z}/p\mathbb{Z}[[q]] \to I_1(N)^h$ over $\mathbb{Z}/p\mathbb{Z}$ such that the triple $(\mathbb{G}_m/q^{\mathbb{Z}}, Id_N, Id_p)$ arises via pull-back from the universal curve. The image of the point $q = 0$ is the cusp ∞ of $I_1(N)$, and q gives a uniformizing parameter in the neighborhood of this cusp. Hence sections of $\underline{\omega}^h$ have a Fourier expansion at ∞, as in (2.6).

The Igusa curve $I_1(N)$ is defined as the smooth compactification of $I_1(N)^h$ over $\mathbb{Z}/p\mathbb{Z}$. The étale covering $\pi\colon I_1(N)^h \to X_1(N)^h$ extends to a ramified covering $\pi\colon I_1(N) \to X_1(N)$, and the line bundle $\underline{\omega}^h$ extends to $I_1(N)$ by $\underline{\omega} = \pi^*(\underline{\omega}_{X_1(N)})$.

PROPOSITION 5.2. *The line bundle $\underline{\omega}$ of $I_1(N)$ has a canonical section a which satisfies*:

1) *a is nonvanishing on $I_1(N)^h$*,
2) *$a^{p-1} = A$ as sections of $\underline{\omega}^{\otimes p-1}$*,
3) *$a(\mathbb{G}_m/q^{\mathbb{Z}}, Id_N, Id_p) = dt/t$, so $a(q) = 1$*,
4) *a has a simple zero at each supersingular point x in $I_1(N) - I_1(N)^h$*,
5) *$a|\langle d\rangle_p = d^{-1}a$ for all $d \in (\mathbb{Z}/p\mathbb{Z})^\times$*.

Proof. Let (E, α, β) be a noncuspidal point of $I_1(N)^h$. There is a unique point P in $E_p^{\acute{e}t} = E_p/\beta(\mu_p)$ such that $f_p(P)\colon E_p \to \mathbb{G}_m$ is the identity map on μ_p (where $f_p\colon E_p \to \operatorname{Hom}(E_p, \mathbb{G}_m)$ is the duality of (1.2)). In other words, for all sections ζ of μ_p we have the formula $e_p(P, \beta(\zeta)) = \zeta$. Let g be a function on E with $\operatorname{div}(g) = p \cdot \{(P) - (0)\}$; the holomorphic differential dg/g in ω_E depends only on β and we define:

$$\alpha(E, \alpha, \beta) = dg/g. \tag{5.3}$$

This gives a non-vanishing section of ω^h on $I_1(N)^h$, once we check that the definition (5.3) extends to the cusps. In fact, the identity $a^{p-1} = A$ shows that a extends to a holomorphic section of $\underline{\omega}$ on $I_1(N)$. To prove it, we recall that the Hasse invariant is given by the formula $A = (C\omega)^p \cdot \omega^{-1}$, where ω is a nonvanishing section of ω_E and C is the Cartier operator. Applying this to $\omega = dg/g$, which satisfies $C(dg/g) = dg/g$, we find $A = a^{p-1}$ over the ordinary points of $I_1(N)$. Since this is an identity between meromorphic sections of $\omega^{\otimes p-1}$ which holds on an open subset of $I_1(N)$, it holds on the entire curve. Since A has a simple zero at each supersingular point of $X_1(N)$, the covering $I_1(N) \to X_1(N)$ is totally ramified at each supersingular point (of degree $p - 1$) and a has a simple zero at each supersingular point x of $I_1(N)$. This proves 1), 2), and 4).

To calculate the Fourier expansion of the section a in a neighborhood of $\infty = (\mathbb{G}_m/q^{\mathbb{Z}}, Id_N, Id_p)$, we remark that in the definition (5.3) we may take $P = q^{-1/p}$ (as we have the formula $e_p(q^{1/p}, \zeta) = \zeta^{-1}$ on the Tate curve) and the function g may be taken to be

$$g(t) = (-t)\frac{\theta(q^{1/p}t)^p}{\theta(t)^p}$$

where $\theta(t) = (1-t)\Pi_{n\geqslant 1}(1-q^n t)(1-q^n/t)$ is the standard Jacobi-Tate theta series. Hence $dg/g = dt/t$ and $a(q) = 1$ as claimed in 3).

By the definition of the action of $\langle d\rangle_p$ on sections of $\underline{\omega}$, we have $a|\langle d\rangle_p(E, \alpha, \beta) = a(E, \alpha, d\beta)$ at ordinary points. But the point P' associated to $\beta' = d\beta$: $\mu_p \hookrightarrow E$ is equal to $d^{-1}P$. Hence $dg'/g' = d^{-1}dg/g$ and $a(E, \alpha, d\beta) = d^{-1}a(E, \alpha, \beta)$. Since the identity $a|\langle d\rangle_p = d^{-1}a$ holds above all ordinary points, it is true on all $I_1(N)$. This proves 5).

Note 5.4. Since the covering $I_1(N) \to X_1(N)$ is totally ramified at the supersingular points, the curve $I_1(N)$ is geometrically connected.

Let S be the affine coordinate ring of $I_1(N)$ over $\mathbb{Z}/p\mathbb{Z}$. Then $S = \bigoplus_\alpha S_\alpha$ is graded by $\alpha \in \mathbb{Z}/(p-1)\mathbb{Z}$, where S_α consists of functions g satisfying $g|\langle d\rangle_p = d^\alpha \cdot g$ for all $d \in (\mathbb{Z}/p\mathbb{Z})^\times$. For any $g \in S$ we let

$$g_\infty = \sum_{n\geqslant 0} a_n q^n$$

be its local expansion in a neighborhood of the cusp ∞ of $I_1(N)$, with respect to the uniformizing parameter q.

PROPOSITION 5.5. *The map $g \mapsto g_\infty$ gives an isomorphism of graded rings $S \xrightarrow{\sim} \tilde{M}$, where $\tilde{M} \subseteq \mathbb{Z}/p\mathbb{Z}[[q]]$ is the ring of Fourier expansions of modular forms for $\Gamma_1(N)$ (mod p).*

The Fourier expansion $\Sigma\, a_n q^n$ in $\tilde{M}_\alpha$ associated to the function $g \in S_\alpha$ has filtration k if and only if $\operatorname{ord}_x(g) \geqslant -k$ *at every supersingular point x of $I_1(N)$, with equality holding for at least one x.*

Proof. We may assume that $g \in S_\alpha$. Let k be an integer with $k \equiv \alpha \pmod{p-1}$ and such that $\operatorname{ord}_x(g) \geqslant -k$ at every supersingular point x of $I_1(N)$. Then by 4) of Proposition 5.2, the element $f = g \cdot a^k$ is a holomorphic section of $\underline{\omega}^{\otimes k}$ on $I_1(N)$. By 5) of Proposition 5.2, this section is fixed by the Galois group of $I_1(N)$ over $X_1(N)$, so f is a modular form of weight k for $\Gamma_1(N)$. Since $a(g) = 1$ by 3) of Proposition 5.2, the Fourier expansion $f(q)$ of f at ∞ is given by $g_\infty = \Sigma\, a_n q^n$. Hence g_∞ is an element of $\tilde{M}_\alpha$.

The map $S_\alpha \to \tilde{M}_\alpha$ is clearly injective. To prove surjectivity we lift an expansion $f(q) \in \tilde{M}_\alpha$ to a modular form $f \in M_k$ and set $g = f/a^k$ in S_α. Then $g_\infty = f(q)$; the filtration of $f(q)$ is the minimal value of k for which a lifting to M_k exists. In this case, the poles of g at supersingular points x have order $\leqslant k$, with equality at one or more x.

Let us also recall the relation between holomorphic differentials on the curve $I_1(N)$ and the spaces of cusp forms M_k^0 of weight $k \leqslant p$ for $\Gamma_1(N)$ (mod p). Recall the isomorphism of line bundles on $X_1(N)$

$$\underline{\omega}^{\otimes 2} \overset{i}{\xrightarrow{\sim}} \Omega^1_{X_1(N)}(cusps)$$

defined in (2.4). Since the map π: $I_1(N) \to X_1(N)$ is totally ramified of degree $e = p - 1$ at the supersingular points, Hurwitz's theorem on computing the canonical bundle for ramified coverings [H, Ch. IV, §2] gives an isomorphism of line bundles on $I_1(N)$:

$$\Omega^1_{I_1(N)} \simeq \pi^*\Omega^1_{X_1(N)}((p-2)\underline{ss}).$$

Here $\underline{ss}$ denotes the divisor of supersingular points on $I_1(N)$, and the multiplicity $p-2=e-1$ occurs as the ramification is tame. Since $\underline{\omega}$ on $I_1(N)$ is simply the pull-back $\pi^*(\underline{\omega})$ from $X_1(N)$, we obtain an isomorphism

$$\underline{\omega}^{\otimes 2}((p-2)\underline{ss}) \simeq \Omega^1_{I_1(N)}(cusps) \tag{5.6}$$

of line bundles on $I_1(N)$. Hence a meromorphic section h of $\omega^{\otimes 2}$ on $I_1(N)$ gives a holomorphic differential on $I_1(N)$ if and only if h is regular at each ordinary point, vanishes at each cusp, and satisfies $\mathrm{ord}_x(h) \geqslant -(p-2)$ at each supersingular point x. If $f \in M_k^0$ with $k \leqslant p$, the section $h = f/a^{k-2}$ satisfies the above conditions, so corresponds to a holomorphic differential ω_f on $I_1(N)$. The Galois group of $I_1(N)$ over $X_1(N)$ acts on ω_f by: $\omega_f|\langle d\rangle_p = d^{k-2}\cdot\omega_f$. Indeed, f is fixed and the group acts on a by d^{-1}, by 5) of Proposition 5.2.

Proposition 5.7. (*Serre, cf.* [S7], [KM; §12.8]) *Assume that* $2 \leqslant k \leqslant p$. *Then the map* $f \mapsto \omega_f = f/a^{k-2}$ *identifies* M_k^0 *with the subspace of holomorphic differentials* $H^0(I_1(N), \Omega^1_{I_1(N)})(k-2)$ *on which the Galois group acts by the character* $\langle d\rangle \mapsto d^{k-2}$. *The expansion of the differential* ω_f *at the cusp* ∞ *is equal to* $f(q)\,dq/q$.

Proof. Since $a(q)=1$, the argument in Proposition 2.8 shows that $\omega_f = f(q)dq/q$ in a neighborhood of ∞. Since the Fourier expansion of $f \in M_k$ determines the form f, this shows that the map $M_k^0 \to H^0(I_1(N), \Omega^1_{I_1(N)})(k-2)$ is an injection. To show it is surjective, we observe that any holomorphic differential v in this eigenspace corresponds, by (5.6), to a meromorphic section h of $\underline{\omega}^{\otimes 2}$. The section h vanishes at the cusps and satisfies $\mathrm{ord}_x(h) \geqslant -(p-2)$ at all supersingular points. In fact, since h is in the $(k-2)$ eigenspace for the Galois action, we must have $\mathrm{ord}_x(h) \equiv -(k-2) \pmod{p-1}$ (or else the section $a^{k-2}h$ would have a fractional order pole at the supersingular point $\pi(x)$ of $X_1(N)$). Hence $\mathrm{ord}_x(h) \geqslant -(k-2)$ at all supersingular points, and $f = a^{k-2}h$ is an element of M_k^0 with $\omega_f = v$.

Arbitrary graded elements of the ring $\tilde{M} = \bigoplus_\alpha \tilde{M}_\alpha$ of Fourier expansions give rise to meromorphic differentials on $I_1(N)$, in the following manner.

Proposition 5.8. *If* $f(q) = \Sigma\, a_n q^n$ *is an element of* $\tilde{M}_\alpha$, *then* $\omega_f = \Sigma\, a_n q^n dq/q$ *is the expansion at the cusp* ∞ *of a meromorphic differential on* $I_1(N)$, *which is regular outside the cusps and supersingular points, and satisfies* $\omega_f|\langle d\rangle_p = d^{\alpha-2}\omega_f$ *for all* $d \in (\mathbb{Z}/p\mathbb{Z})^\times$.

At the cusps the poles of ω_f *have order* $\leqslant 1$, *and* ω_f *is regular at all cusps if and only if* $f(q)$ *is the expansion of a cusp form. At each supersingular point* x *we have* $\mathrm{ord}_x(\omega_f) \equiv (p-\alpha) \pmod{p-1}$. *The expansion* $f(q)$ *has filtration* k *if and only if* $\mathrm{ord}_x(\omega_f) \geqslant (p-k)$ *at all supersingular points* x, *with equality holding for at least one* x.

If $h(q) \in \tilde{M}_\alpha$ corresponds to the function $g \in S_\alpha$, then $f(q) = \theta h(q) \in \tilde{M}_{\alpha+2}$ corresponds to the exact differential $\omega_f = dg$ on $I_1(N)$.

Proof. Assume $f(q)$ has filtration $k \equiv \alpha$. Then ω_f is the expansion at ∞ of the meromorphic section f/a^{k-2} of $\underline{\omega}^{\otimes 2} \simeq \Omega^1_{I_1(N)}(cusps + 2 - p)\underline{ss})$. As a section of $\omega^{\otimes 2}$ it is regular at the cusps and satisfies $\operatorname{ord}_x(f/a^{k-2}) \geqslant (2 - k)$ at each supersingular point, with equality holding for at least one x. The corresponding statements for ω_f as a meromorphic section of $\Omega^1_{I_1(N)}$ follow immediately.

The fact that θh corresponds to dg follows from a comparison of expansions at ∞.

D. Ulmer has remarked that the differential $\omega_f = dq/q$ associated to the modular function $1 \in M_0$ classifies the extension of group schemes

$$0 \to \mu_p \xrightarrow{\beta} E_p \to \mathbb{Z}/p\mathbb{Z} \to 0$$

over $I_1(N)^h - cusps$. We also mention the fact (implicit in 5.7 and 5.8) that the automorphism $\langle d \rangle_p$ acts by multiplication by d^{-1} on the tangent space of each supersingular point x on $I_1(N)$. (The action is defined by transport of structure.)

The operators $\langle d \rangle_N$, for $d \in (\mathbb{Z}/N\mathbb{Z})^\times$, T_l, and U_l (for $l \neq p$) define correspondences on the curve $I_1(N)$ over $\mathbb{Z}/p\mathbb{Z}$, using the formulae (3.1), (3.11), and (3.13). Hence they define endomorphisms of the space of holomorphic differential forms on $I_1(N)$, which preserve the eigencomponents for the action of the Galois group $(\mathbb{Z}/p\mathbb{Z})^\times$.

PROPOSITION 5.9. *The isomorphism $M_k^0 \rightleftarrows H^0(I_1(N), \Omega^1_{I_1(N)})(k-2)$ of Proposition 5.7 commutes with the action of $\langle d \rangle_N$, T_l, and U_l on differentials and modular forms. It transforms the endomorphism U_p of M_k^0 to the endomorphism $Ver_p = (Fr_p)^t$ of the holomorphic differentials.*

Proof. The claim for $\langle d \rangle_N$ is clear, as $\omega_{f|\langle d \rangle_N}(E, \alpha, \beta) = h|\langle d \rangle_N(E, \alpha, \beta)$ where $h = f/a^{k-2}$ is the corresponding meromorphic section of $\underline{\omega}^{\otimes 2}$. But $h|\langle d \rangle_N = f|\langle d \rangle_N/a^{k-2}$ as the section a of $\underline{\omega}$ does not depend on α, so is fixed by $\langle d \rangle_N$. Hence $\omega_{f|\langle d \rangle_N} = \omega_f|\langle d \rangle_N$.

Now suppose $f \in M_k^0$ has Fourier expansion $\Sigma_{n \geqslant 1} a_n q^n$, and for $l \nmid Np$, $f|\langle l \rangle_N$ has expansion $\Sigma_{n \geqslant 1} b_n q^n$. Then $F|T_l(q) = \Sigma a_{nl} q^n + l^{k-1} \Sigma b_n q^{nl}$. On the other hand, $h|T_l(q) = \Sigma a_{nl} q^n + \varepsilon_p(l) \cdot l \Sigma b_n q^{nl}$, as h has weight 2 and character ε_p for the group $(\mathbb{Z}/p\mathbb{Z})^\times$. Since $\varepsilon_p(l) = l^{k-2}$, we see that $\omega_{f|T_l} = \omega_f|T_l$ as claimed. A similar argument, using Fourier expansions at ∞, works for U_l and U_p.

When $k = p + 1$, the argument of Propositions 5.7 and 5.9 gives the following (whose proof we omit).

PROPOSITION 5.10. *The map $f \mapsto \omega_f = f/A$ identifies M_{p+1}^0 with the space $H^0(X_1(N), \Omega^1_{X_1(N)}(\underline{ss}))$ of differentials of the third kind on $X_1(N)$ over $\mathbb{Z}/p\mathbb{Z}$ with simple poles at the supersingular points. This isomorphism commutes with the action of the Hecke operators, and the expansion of ω_f at the cusp ∞ is equal to $f(q)dq/q$. The differential ω_f is holomorphic on $X_1(N)$ if and only if $f \in AM_2^0$.*

§6. The curve $X_1(Np)$ over $\mathbb{Z}[1/Np][\zeta_p]$. We now consider the curve $X_1(Np)$ in more detail, where $N \geqslant 4$ and p is prime to N. By the results of §2, this curve is smooth and proper over $\mathbb{Z}[1/Np]$. We now consider certain automorphisms w_ζ of $X_1(Np)$, associated to primitive p^{th} roots of unity ζ, which are rational over the étale extension $\mathbb{Z}[1/Np][\zeta] = \mathbb{Z}[1/Np][\zeta_p]$. These automorphisms were studied, in the case where $N = 1$, by Mazur and Tate [MT, §5] and Wiles [Wi].

The noncuspidal points of $X_1(Np)$ correspond to triples (E, α, β), where E is an elliptic curve over a $\mathbb{Z}[1/Np]$-algebra and α and β are embeddings of group schemes $\alpha\colon \mu_N \hookrightarrow E_N$, $\beta\colon \mu_p \hookrightarrow E_p$. (We may separate α and β as p is prime to N.) For $d \in (\mathbb{Z}/N\mathbb{Z})^\times$ we have the automorphism

$$(6.1) \qquad \langle d\rangle_N(E, \alpha, \beta) = (E, d\alpha, \beta),$$

and for $d \in (\mathbb{Z}/p\mathbb{Z})^\times$ we have the automorphism

$$(6.2) \qquad \langle d\rangle_p(E, \alpha, \beta) = (E, \alpha, d\beta).$$

If d is prime to Np, and $\langle d\rangle$ is the automorphism of $X_1(Np)$ defined in (3.1), we have $\langle d\rangle = \langle d\rangle_N \cdot \langle d\rangle_p$. Similarly, we have defined Hecke correspondences T_l (for $l \nmid Np$), U_l (for $l|N$), and U_p on $X_1(Np)$ over $\mathbb{Z}[1/Np]$.

Now let ζ be a primitive p^{th} root of unity: this gives an isomorphism

$$(6.3) \qquad \begin{aligned} i_\zeta\colon \mu_p &\xrightarrow{\sim} \mathbb{Z}/p\mathbb{Z} \\ \zeta &\mapsto 1 \end{aligned}$$

of finite groups schemes over $\mathbb{Z}[1/Np][\zeta_p]$. Let (E, α, β) be a noncuspidal point of $X_1(Np)$ over $\mathbb{Z}[1/Np][\zeta_p]$. Let $E' = E/\beta(\mu_p)$ be the p-isogenous curve, and $\varphi\colon E \to E'$ the corresponding p-isogeny. Let $\alpha' = \varphi\alpha\colon \mu_N \hookrightarrow E'$; this is an embedding as N is prime to p. If $e_p\colon E_p \times E_p \to \mu_p$ is the pairing defined in (1.1), there is a unique class P_β in $E_p/\beta(\mu_p)$ such that $e_p(\beta(z), P_\beta) = z$ for all $z \in \mu_p$. Let $P'_\beta = \varphi(P_\beta)$, which is well defined in E'_p. Finally, let $\beta'\colon \mu_p \hookrightarrow E'_p$ be the embedding of group schemes defined by $\beta'(\zeta) = P'_\beta$. We define

$$(6.4) \qquad w_\zeta(E, \alpha, \beta) = (E', \alpha', \beta')$$

as an automorphism of $X_1(Np)$ over $\mathbb{Z}[1/Np][\zeta_p]$.

Note that the choice of ζ appears only in the final definition of embedding β'. If we let $b'\colon \mathbb{Z}/p\mathbb{Z} \hookrightarrow E'_p$ be the embedding of group schemes with $b(1) = P'_\beta$, then the map

$$(6.5) \qquad v(E, \alpha, \beta) = (E', \alpha', b')$$

gives an isomorphism from $X_1(Np)$ to the curve $X_1(Np)^*$ which classifies triples (E, α, b) where $b\colon \mathbb{Z}/p\mathbb{Z} \hookrightarrow E_p$. The choice of ζ gives an isomorphism from $X_1(Np)^*$

to $X_1(Np)$:

$$u_\zeta(E', \alpha', b') = (E', \alpha', b' \circ i_\zeta) \tag{6.6}$$

and we have $w_\zeta = u_\zeta \circ v$.

PROPOSITION 6.7. 1) $w_\zeta(\mathbb{G}_m/q^{\mathbb{Z}}, Id_N, Id_p) = (\mathbb{G}_m/q^{p\mathbb{Z}}, p\, Id_N, \zeta \mapsto q)$.

2) *For* $d \in (\mathbb{Z}/p\mathbb{Z})^\times$, $w_{\zeta^d} = \langle d \rangle_p^{-1} w_\zeta = w_\zeta \cdot \langle d \rangle_p$.

3) *Let* $w = w_\zeta$ *for any primitive* p^{th} *root of unity* ζ. *Then*

$$w^2 = \langle p \rangle_N \cdot \langle -1 \rangle_p$$

$$w \cdot \langle d \rangle_N = \langle d \rangle_N \cdot w \quad \textit{for all } d \in (\mathbb{Z}/N\mathbb{Z})^\times$$

$$w \cdot T_l = \langle l \rangle_p^{-1} T_l \cdot w \quad \textit{for all } l \nmid Np$$

$$w \cdot U_l = \langle l \rangle_p^{-1} U_l \cdot w \quad \textit{for all } l | N.$$

Proof. 1) We calculate over $\mathbb{Z}[1/Np][\zeta_p]((q))$, where $\mathbb{G}_m/q^{\mathbb{Z}} = E$ is an elliptic curve. The isogeny

$$\varphi\colon \mathbb{G}_m/q^{\mathbb{Z}} \xrightarrow[p]{} \mathbb{G}_m/q^{p\mathbb{Z}}$$

has kernel $\beta(\mu_p) = Id_p(\mu_p)$, so $E' = \mathbb{G}_m/q^{p\mathbb{Z}}$ and $\alpha' = \varphi\alpha = p \cdot Id_N$. The point P_β is equal to $q^{1/p}$, by formula (1.8), so $P'_\beta \equiv q$ in $\mathbb{G}_m/q^{p\mathbb{Z}}$. Hence $\beta'\colon \mu_p \hookrightarrow E'_p$ is the map taking ζ to q.

2) Since $i_{\zeta^d} = i_\zeta \circ \langle d \rangle_p = \langle d \rangle_p^{-1} i_\zeta$, this follows from the formula: $w_{\zeta^d} = u_{\zeta^d} \circ v$.

3) Let $w^2(E, \alpha, \beta) = (E'', \alpha'', \beta'')$. Since ${}^t\varphi\colon E' \to E$ is the p-isogeny with kernel $\beta'(\mu_p)$, we have $E'' = E$. Since ${}^t\varphi \circ \varphi = p$, we have $\alpha'' = p \cdot \alpha$. As the Weil pairing e_p is alternating, $\beta'' = -\beta$. Hence $w^2 = \langle p \rangle_N \cdot \langle -1 \rangle_p$.

The fact that w commutes with $\langle d \rangle_N$ is clear. To derive the commutation laws with T_l and U_l, we use the identity

$$e_p(\psi\beta(z), \psi P_\beta) = z^{\deg \psi}$$

where $\psi\colon E \to F$ is any isogeny of degree prime to p. Hence $\psi P_\beta = \langle \deg \psi \rangle_p \cdot P_{\psi_\beta}$ in $F_p/\psi\beta(\mu_p)$. The rest is a consequence of the definitions of T_l and U_l, and we leave the proof to the reader.

The automorphism w_ζ of $X_1(Np)$ acts on the space of holomorphic differentials, by pull-back. Via the isomorphism (2.4) this induces an action of w_ζ on the space of cusp forms of weight 2 for $\Gamma_1(Np)$ over $\mathbb{Z}[1/Np][\zeta_p]$. One can verify, using the method of Proposition 3.15, that the action is given by

$$f | w_\zeta (E, \alpha, \beta) = \frac{1}{p} \varphi^*(f(E', \alpha', \beta')) \tag{6.8}$$

where (E', α', β') is defined in (6.4), $\varphi\colon E \to E'$ is the associated p-isogeny, and $\varphi^*\colon \omega_{E'}^{\otimes 2} \to \omega_E^{\otimes 2}$ the pull-back.

If we conjugate the correspondence U_p by the automorphism w_ζ, we obtain a new correspondence (cf. [S7], [Wi, §2]):

$$U_p' = w_\zeta^{-1} U_p w_\zeta \tag{6.9}$$

on $X_1(Np)$ over $\mathbb{Z}[1/Np][\zeta_p]$. By part 2) of Proposition 6.7, the correspondence U_p' is independent of the choice of primitive p^{th} root of unity ζ (as U_p commutes with $\langle d\rangle_p$). Hence U_p' is defined over $\mathbb{Z}[1/Np]$. Using (6.8), we get an action of U_p' on the space of modular forms of weight 2 for $\Gamma_1(Np)$ over $\mathbb{Z}[1/Np]$, and this action is compatible with its natural action as a correspondence on holomorphic differentials on $X_1(Np)$. By part 3) the operator U_p' commutes with the action of the operators $\langle d\rangle_N$, $\langle d\rangle_p$, T_l and U_l on the space of forms of weight 2. It usually does not commute with U_p.

We now determine the action of U_p' on Fourier expansions.

Proposition 6.10. *Let f be a modular form of weight 2 for $\Gamma_1(Np)$. Then*

$$f\,|\,U_p'(q) = \sum_{d\in(\mathbb{Z}/p\mathbb{Z})^*} f\,|\,w_{\zeta^d}(q\zeta^d) + p\cdot f\,|\,\langle p\rangle_N(q^p).$$

Proof. We recall that, by definition of the Fourier expansion in (2.6):

$$f\,|\,U_p'(\mathbb{G}_m/q^{\mathbb{Z}}, Id_N, Id_p) = f\,|\,U_p'(q)\cdot(dt/t)^{\otimes 2}.$$

Since $w_\zeta(\mathbb{G}_m/q^{\mathbb{Z}}, Id_N, Id_p) = (\mathbb{G}_m/q^{p\mathbb{Z}}, pId_N, \zeta \mapsto q)$ and the p-isogeny is $\varphi\colon \mathbb{G}_m/q^{\mathbb{Z}} \xrightarrow[p]{} \mathbb{G}_m/q^{p\mathbb{Z}}$, we have

$$\begin{aligned} f\,|\,U_p'(\mathbb{G}_m/q^{\mathbb{Z}}, Id_N, Id_p) &= f\,|\,w_\zeta^{-1} U_p w_\zeta(\mathbb{G}_m/q^{\mathbb{Z}}, Id_N, Id_p) \\ &= \frac{1}{p}\varphi^*(f\,|\,w_\zeta^{-1} U_p(\mathbb{G}_m/q^{p\mathbb{Z}}, pId_N, \zeta\mapsto q)). \end{aligned}$$

By the definition of U_p in (3.13), we have

$$\begin{aligned} &U_p(\mathbb{G}_m/q^{p\mathbb{Z}}, pId_N, \zeta\mapsto q) \\ &\quad= \sum_{d\in(\mathbb{Z}/p\mathbb{Z})^*} (\mathbb{G}_m/(q\zeta^d)^{\mathbb{Z}}, pId_N, \zeta\mapsto\zeta^{-d}) + (\mathbb{G}_m/q^{p^2\mathbb{Z}}, p^2Id_N, \zeta\mapsto q^p) \end{aligned}$$

as the isogenies $\psi_d\colon \mathbb{G}_m/q^{p\mathbb{Z}} \xrightarrow{1} \mathbb{G}_m/(q\zeta^d)^{\mathbb{Z}}$ for $d \in (\mathbb{Z}/p\mathbb{Z})^\times$ and $\psi\colon \mathbb{G}_m/q^{p\mathbb{Z}} \xrightarrow{p} \mathbb{G}_m/q^{p^2\mathbb{Z}}$ are those isogenies of degree p whose kernel does not meet the subgroup $\langle q\rangle$. Letting

$g = f|w_\zeta^{-1}$, we have

$$(6.11)\quad f|U_p'(\mathbb{G}_m/q^{\mathbb{Z}}, Id_N, Id_p) = \frac{1}{p^2}\varphi^* \sum_{d\in(\mathbb{Z}/p\mathbb{Z})^\times} \psi_d^*(g(\mathbb{G}_m/(q\zeta^d)^{\mathbb{Z}}, pId_N, \zeta\mapsto\zeta^{-d}))$$

$$+\frac{1}{p^2}\varphi^*\psi^* g(\mathbb{G}_m/q^{p^2\mathbb{Z}}, p^2Id_N, \zeta\mapsto q^p).$$

But $w_\zeta^2 = \langle p\rangle_N\cdot\langle -1\rangle_p$, so $g = f|w_\zeta^{-1} = f|w_\zeta\cdot\langle p\rangle_N^{-1}\cdot\langle -1\rangle_p^{-1}$. Hence

$$\begin{aligned} &g(\mathbb{G}_m/(q\zeta^d)^{\mathbb{Z}}, pId_N, \zeta\mapsto\zeta^{-d}) \\ &= f|w_\zeta(\mathbb{G}_m/(q\zeta^d)^{\mathbb{Z}}, Id_N, \zeta\mapsto\zeta^d) \\ &= f|w_\zeta\langle d\rangle_p(\mathbb{G}_m/(q\zeta^d)^{\mathbb{Z}}, Id_N, Id_p) \\ &= f|w_{\zeta^d}(\mathbb{G}_m/(q\zeta^d)^{\mathbb{Z}}, Id_N, Id_p). \end{aligned}$$

Since $\varphi^*((dt/t)^{\otimes 2}) = p^2(dt/t)^{\otimes 2}$ and $\psi_d^*((dt/t)^{\otimes 2}) = (dt/t)^{\otimes 2}$, the first sum in (6.11) gives a contribution of

$$\sum_{d\in(\mathbb{Z}/p\mathbb{Z})^*} f|w_{\zeta^d}(q\zeta^d)$$

to the Fourier expansion of $f|U_p'$. The second term in (6.11) corresponds to

$$\begin{aligned} g(\mathbb{G}_m/q^{p^2\mathbb{Z}}, p^2Id_N, \zeta\mapsto q^p) &= f|w_\zeta(\mathbb{G}_m/q^{p^2\mathbb{Z}}, pId_N, \zeta\mapsto q^{-p}) \\ &= \frac{1}{p}\rho^* f(\mathbb{G}_m/q^{p\mathbb{Z}}, pId_N, \zeta\mapsto\zeta) \end{aligned}$$

as the relevant p-isogeny for computing w_ζ is $\rho\colon \mathbb{G}_m/q^{p^2\mathbb{Z}} \underset{i}{\to} \mathbb{G}_m/q^{p\mathbb{Z}}$, and $P = \zeta$ by (1.8). Since $\psi^*((dt/t)^{\otimes 2}) = p^2(dt/t)^{\otimes 2}$ and $\rho^*(dt/t)^{\otimes 2} = (dt/t)^{\otimes 2}$, the second term contributes $p\cdot f|\langle p\rangle(q^p)$ to the Fourier expression of $f|U_p'$. This completes the proof.

PROPOSITION 6.12. *Let f be a modular form of weight 2 for $\Gamma_1(Np)$ which satisfies $\Sigma_d f|\langle d\rangle_p = 0$. Then $f|U_p'U_p = p\cdot f|\langle p\rangle_N$.*

Proof. Let $\Sigma\, a_n q^n$ be the Fourier expansion of $f|\langle p\rangle_N$. Then by the previous proposition

$$f|U_p'(q) = \sum_{d\in(\mathbb{Z}/p\mathbb{Z})^\times} f|w_{\zeta^d}(q\zeta^d) + p\Sigma\, a_n q^{np}.$$

The coefficient of q^{np} in the sum is equal to zero. Indeed, this is equal to the coefficient of q^{np} in $\Sigma_d f|w_{\zeta^d}(q)$ as $(\zeta^d)^p = 1$ for all p. But $\Sigma_d f|w_{\zeta^d} = (\Sigma_d f|\langle d^{-1}\rangle)|w_\zeta$ and $\Sigma_d f|\langle d^{-1}\rangle = 0$ by hypothesis.

Hence the coefficients of q^{np} in $f|U_p'(q)$ is equal to pa_n, and $f|U_p'U_p(q) = p\Sigma a_n q^n = p\cdot f|\langle p\rangle_N(q)$. Therefore $f|U_p'U_p = p\cdot f|\langle p\rangle_N$ by the q-expansion principle.

The space of modular forms of weight 2 for $\Gamma_1(Np)$ over $\mathbb{Q}$ is the direct sum of subspaces $M' \oplus M''$, where M' consists of forms f satisfying $f|\langle d\rangle_p = f$ for all $d \in (\mathbb{Z}/p\mathbb{Z})^\times$ and M'' consists of forms f satisfying $\Sigma_d f|\langle d\rangle_p = 0$. Both subspaces are stable under the operators $\langle d\rangle_N$, T_l, U_l, U_p, and U_p', as these operators commute with the $\langle d\rangle_p$.

Corollary 6.13. *On the subspace M'' the operators U_p and U_p' commute. They are invertible, and satisfy $U_p'U_p = p\cdot\langle p\rangle_N$.*

Proof. The relation $U_p'U_p = p\langle p\rangle_N$ on M'' follows from Proposition 6.12. Hence U_p' and U_p are invertible. Since $p\cdot\langle p\rangle_N$ commutes with U_p, so does $U_p' = p\cdot\langle p\rangle_N\cdot U_p^{-1}$.

Note. On the subspace of M' consisting of forms which are *new* at p, the operators U_p and U_p' also commute. In fact, one has $U_p' = U_p = -w$ on M'_{new}, so $U_p'U_p = \langle p\rangle_N$ (cf. [Li]). The failure of U_p and U_p' to commute on M is therefore due to the presence of forms in M' which are old at p.

Proposition 6.14. *Let $F = \Sigma A_n q^n$ be a newform ($=$normalized cuspidal eigenform) of weight 2 and character $\varepsilon = \varepsilon_N\cdot\varepsilon_p$ for $\Gamma_1(Np)$ with coefficients in an extension of $\mathbb{Q}$. Then $F|w_\zeta = c_\zeta\cdot F'$, where $F' = \Sigma A_n' q^n$ is a newform of weight 2 and character $\varepsilon' = \varepsilon_N\cdot\varepsilon_p^{-1}$ for $\Gamma_1(Np)$ and c_ζ is a nonzero constant. We have*

$$A_n' = A_n/\varepsilon_p(n) \quad \text{for all } n \text{ prime to } p. \tag{6.15}$$

If $\varepsilon_p = 1$ then $F' = F$, $A_p^2 = \varepsilon_N(p)$, and $c_\zeta = -A_p$. If $\varepsilon_p \neq 1$ then

$$A_p'\cdot A_p = p\varepsilon_N(p) \tag{6.16}$$

and the constant c_ζ is given by the formula:

$$c_\zeta = \frac{\varepsilon_p(-1)\varepsilon_N(p)\sum_d \varepsilon_p(d)\zeta^d}{A_p} \tag{6.17}$$

Proof. By Proposition 6.7 it follows that $F|w_\zeta$ is a form with character $\varepsilon' = \varepsilon_N\cdot\varepsilon_p^{-1}$ which satisfies

$$(F|w_\zeta)|T_l = A_l/\varepsilon_p(l)\cdot(F|w_\zeta)$$

$$(F|w_\zeta)|U_l = A_l/\varepsilon_p(l)\cdot(F|w_\zeta)$$

for $l \neq p$. To show $F|w_\zeta = c_\zeta\cdot F'$, where F' is a newform, it suffices to show $F|w_\zeta$ is an eigenvector for U_p. We will prove this when $\varepsilon_p \neq 1$; a proof when $\varepsilon_p = 1$ (assuming

F is new at p) is given in [ALi], which also gives the formulae for A'_p and c_ζ in this case.

If $\varepsilon_p \neq 1$ then F lies in the extension of scalars of the subspace M'', as $\Sigma_d F|\langle d\rangle_p = \Sigma\, \varepsilon_p(d)\cdot F = 0$. Hence

$$\begin{aligned}(F|w_\zeta)|U_p &= F|U'_p|w_\zeta\\ &= \frac{p\cdot \varepsilon_N(p)}{A_p}\cdot F|w_\zeta,\end{aligned}$$

as $U'_p = p\langle p\rangle_N U_p^{-1}$ on M''. This shows $F|w_\zeta = c_\zeta F'$, where F' has p^{th} coefficient $A'_p = p\varepsilon_N(p)/A_p$. The constant c_ζ is equal to the coefficient of q in $F|w_\zeta(q)$, which we may determine as follows.

First, observe that $F|w_{\zeta^d} = F|\langle d^{-1}\rangle w_\zeta = \varepsilon_p(d^{-1})\cdot F|w_\zeta$ by Prop. 6.10. By Proposition 6.12, the coefficient of q in $F|U'_p(q)$ is equal to:

$$c_\zeta\cdot \sum_d \varepsilon_p(d^{-1})\zeta^d.$$

But by Corollary 6.13, this coefficient is equal to $p\varepsilon_N(p)/A_p$. Hence

$$c_\zeta = \frac{p}{\sum \varepsilon_p(d^{-1})\zeta^d}\cdot\frac{\varepsilon_N(p)}{A_p}.$$

Since $(\Sigma\, \varepsilon_p(d)\zeta^d)\cdot(\Sigma\, \varepsilon_p(d^{-1})\zeta^d) = p\cdot\varepsilon_p(-1)$ [L, Ch. I], we obtain formula (6.17).

§7. A model for $X_1(Np)$ over $\mathbb{Z}[1/N][\zeta_p]$. We now describe a stable model X for the curve $X_1(Np)$ over the base $\mathbb{Z}[1/N][\zeta_p]$, which was introduced by Deligne and Rapoport [DR, V, §2]. To be completely accurate, we note that the scheme $X' = \mathcal{M}_{\Gamma_{00}(N)\cap\Gamma'_{00}(p)}$ defined by Deligne and Rapoport is actually a model for the curve $X_1(Np)'$, which classifies triples (E, a, β) with a: $\mathbb{Z}/N\mathbb{Z} \hookrightarrow E_N$. However, the schemes X and X' become isomorphic over $\mathbb{Z}[1/N][\zeta_p, \zeta_N]$, and one can obtain X over the base $\mathbb{Z}[1/N][\zeta_p]$ by Galois descent.

To define X, we let $\mathcal{X}(Np)$ be the projective scheme over $\mathbb{Z}[1/N]$ which represents the functor of generalized elliptic curves E together with a "Drinfeld basis" of E_{Np} [KM, 5.1.1]. This is a regular scheme of dimension 2, with a natural action of the group $GL_2(\mathbb{Z}/N\mathbb{Z}) \times GL_2(\mathbb{Z}/p\mathbb{Z})$. We let X be the quotient of this scheme by the finite group $H_N \times H_p$, where $H_N = \left\{\begin{pmatrix} a & b\\ c & d\end{pmatrix} \in GL_2(\mathbb{Z}/N\mathbb{Z})\text{: } c \equiv 0\ (N) \text{ and } d \equiv 1 \pmod N\right\}$ and $H_p = \left\{\begin{pmatrix} 1 & b\\ 0 & 1\end{pmatrix} \in GL_2(\mathbb{Z}/p\mathbb{Z})\right\}$. This quotient is defined over $\mathbb{Z}[1/N][\zeta_p]$ and has the following properties.

Proposition 7.1. 1) *X is a regular scheme of dimension 2, and the morphism $X \to \operatorname{Spec} \mathbb{Z}[1/N][\zeta_p]$ has only ordinary double points as singularities.*

2) *X is smooth over Spec $\mathbb{Z}[1/Np][\zeta_p]$. The special fibre X_0 of X in characteristic p is a reduced curve over $\mathbb{Z}/p\mathbb{Z}$, consisting of two smooth projective curves I and I' crossing transversally at a finite set Σ of double points.*

3) *The component I is canonically isomorphic to the Igusa curve $I_1(N)$, and Σ is the set of supersingular points on $I_1(N)$. The component I' is the smooth compactification of the moduli of triples (E, α, b), where b: $\mathbb{Z}/p\mathbb{Z} \hookrightarrow E_p$.*

Proof. This is a restatement of Théorème 2.12 in [DR, V, §2].

The sections of X which do not meet Σ correspond to triples (E, α, β) or to triples (E, α, b) over $\mathbb{Z}[1/N][\zeta_p]$-algebras, where β: $\mu_p \hookrightarrow E_p$ and b: $\mathbb{Z}/p\mathbb{Z} \hookrightarrow E_p$ are embeddings of group schemes. The sections of the type (E, α, β) meet the component I and those of the type (E, α, b) meet the component I'. The map v defined in (6.5) gives a canonical automorphism of X over $\mathbb{Z}[1/N][\zeta_p]$ which interchanges these two types of sections and induces an isomorphism $I \overset{\sim}{\rightarrow} I'$ in characteristic p.

Fix a primitive p^{th} root of unity ζ, and let i_ζ: $\mu_p \overset{\sim}{\rightarrow} \mathbb{Z}/p\mathbb{Z}$ be defined as in (6.3). We then have an isomorphism from X to $X_1(Np)$ over the base $\mathbb{Z}[1/N][\zeta_p]$. The sections (E, α, β) of X are mapped to the corresponding points of $X_1(Np)$, and the sections (E, α, b) are mapped to the points $(E, \alpha, \beta = b \circ i_\zeta)$. The automorphism v of X then induces the automorphism w_ζ of $X_1(Np)$. We henceforth identify $X_1(Np)$ with the "general fibre" of X over $\mathbb{Z}[1/N][\zeta]$.

The automorphisms $\langle d\rangle_N$ and $\langle d\rangle_p$ of $X_1(Np)$ extend to X, and induce the automorphisms with the same name on the component $I_1(N) = I$ of X_0. Since the generic fibre $X_1(Np)$ of X can be defined over $\mathbb{Z}[1/Np]$, there is a semi-linear action of the Galois group Γ of the covering $\mathbb{Z}[1/Np][\zeta_p]$ over $\mathbb{Z}[1/Np]$ on X. Since this covering is totally ramified at p, this action induces a geometric action of Γ on the special fibre X_0 (cf. [ST, pg. 483]). We may identify Γ with $(\mathbb{Z}/p\mathbb{Z})^\times$ by letting σ_d be the automorphism with $\sigma_d(\zeta_p) = \zeta_p^d$.

PROPOSITION 7.2. *The element σ_d of Γ acts on X_0 via the automorphism $1 \times \langle d\rangle_p^{-1}$ of $I \times I'$.*

Proof. For sections of X meeting $I - \Sigma$, we have the formula $\sigma(E, \alpha, \beta) = (E^\sigma, \alpha^\sigma, \beta^\sigma)$ for any $\sigma \in \Gamma$. Since Γ acts trivially on the residue field at p, this gives the trivial automorphism of I.

For sections of X meeting $I' - \Sigma$, we have the formula

$$\sigma_d(E, \alpha, b) = (E^{\sigma_d}, \alpha^{\sigma_d}, d^{-1} \cdot b^{\sigma_d}). \tag{7.3}$$

Indeed, (E, α, b) corresponds to the point $(E, \alpha, \beta = b \circ i_\zeta)$ in the general fibre, so its conjugate by $\sigma \in \Gamma$ is the point $(E^\sigma, \alpha^\sigma, \beta^\sigma)$. Write $\beta^\sigma = b' \circ i_\zeta$; for $\sigma = \sigma_d$ we must show that $b' = d^{-1} \cdot b^{\sigma_d}$. But

$$\begin{aligned} b'(d) = \beta^\sigma(\zeta^d) &= \beta^\sigma(\zeta^\sigma) \quad \text{if } \sigma = \sigma_d \\ &= \beta(\zeta)^\sigma \\ &= b(1)^\sigma. \end{aligned}$$

This proves (7.3); since Γ acts trivially on the residue field at p we have $\sigma_d = \langle d \rangle_p^{-1}$ on I'. (We note that the action of Γ on X_0 is independent of the choice of ζ used to define it.)

The correspondences T_l and U_l of $X_1(Np)$ (for $l \neq p$) preserve the sections reducing to $I - \Sigma$ or $I' - \Sigma$ (modulo p) on X, and their induced action on $I = I_1(N)$ is the one described in §5. The correspondence U_p preserves the sections reducing to $I - \Sigma$, and induces the correspondence $U_p = Ver_p$ on $I = I_1(N)$. The correspondence U_p does *not* preserve the sections meeting $I' - \Sigma$, and hence the correspondence U_p' does not preserve the sections meeting I. From Proposition 6.10 one can derive the formula (cf. [Wi, §5]):

$$(7.4) \qquad U_p' \equiv \sum_{d \in (\mathbb{Z}/p\mathbb{Z})^*} w_{\zeta^d} + \langle p \rangle_N Fr_p \quad \text{on } I = I_1(N).$$

The automorphism $w = w_\zeta$ preserves the set Σ of supersingular points on $I_1(N)$, and its action on Σ is independent of the choice of ζ (as $\langle d \rangle_p$ fixes each supersingular point). If $x \in \Sigma$ covers the supersingular point (E, α) on $X_1(N)$, then $w(x)$ covers the supersingular point (E', α'), where $E' = E^{(p)}$, $\varphi = Fr_p \colon E \to E^{(p)}$ is the associated p-isogeny, and $\alpha' = \varphi\alpha$. Hence $\alpha'(z) = \alpha(z)^p = \alpha^{(p)}(z^p)$ for all $z \in \mu_N$, and $w(x)$ covers the point $Fr_p(E, p\alpha)$ of $X_1(N)$. Consequently:

$$(7.5) \qquad \begin{cases} w = Fr_p \cdot \langle p \rangle_N = \langle p \rangle_N \cdot Fr_p \\ U_p' = p \cdot \langle p \rangle_N \cdot Fr_p \end{cases} \quad \text{on } \Sigma.$$

§8. Regular differentials. Let ζ be a primitive p^{th} root of unity in an algebraic closure of $\mathbb{Q}_p$. We now consider the scheme X over the base $\mathcal{O} = \mathbb{Z}_p[\zeta]$, and identify its general fibre with $X_1(Np)$ over the field $K = \mathbb{Q}_p(\zeta)$ (using the isomorphism i_ζ as in §7). The element $\pi = 1 - \zeta$ is a uniformizing parameter for the discrete valuation ring $\mathcal{O}$, and $\mathcal{O}/\pi\mathcal{O} \xrightarrow{\sim} \mathbb{Z}/p\mathbb{Z}$.

Since $X \to Spec\ \mathcal{O}$ is a morphism which is purely of dimension 1 and is locally a complete intersection, the dualizing sheaf $\Omega_{X/\mathcal{O}}$ of "regular differentials" is invertible on X [DR, I.2: in this reference the dualizing sheaf is denoted by $\omega_{X/\mathcal{O}}$]. Let $L = H^0(X, \Omega_{X/\mathcal{O}})$; we call elements of the finitely generated $\mathcal{O}$-module L regular differentials on X.

If R is flat over $\mathcal{O}$, pull-back gives an isomorphism

$$L \otimes_{\mathcal{O}} R \xrightarrow{\sim} H^0(X/R, \Omega_{X/R}).$$

In particular, this holds for $R = K$ where $X = X_1(Np)$ is smooth and $\Omega_{X/K} \simeq \Omega^1_{X_1(Np)/K}$. Hence elements of $L \otimes K$ correspond to holomorphic differentials on $X_1(Np)$ over K, and hence to cusp forms of weight 2 for $\Gamma_1(Np)$ over K. We again emphasize that this identification requires the choice of a p^{th} root of unity ζ.

In fact, L is torsion-free and defines an $\mathcal{O}$-lattice in the space of cusp forms of weight 2. This follows from the argument in [M, II, §3], which also shows that pull-back gives an isomorphism

$$(8.2) \qquad L/\pi L \stackrel{\sim}{\to} H^0(X_0, \Omega_{X_0/(\mathbb{Z}/p\mathbb{Z})}) = L_0 .$$

The right hand side consists of the regular differentials on X_0, the special fibre. This space can be identified with a subspace of the differentials of the third kind on the normalization $\tilde{X}_0 = I \amalg I'$ [S1, Ch. IV, §3]: a pair (v, v') of differentials of the third kind on $I \amalg I'$ corresponds to a regular differential ω on X_0 if and only if v and v' are regular outside $\Sigma = I \cap I'$ and satisfy

$$(8.3) \qquad Res_x(v) + Res_x(v') = 0$$

for all $x \in \Sigma$. This follows from the fact that X_0 has only ordinary double points as singularities.

For $\omega \in L$ we write $\omega \equiv (v, v') \pmod{\pi L}$ if v and v' are the differentials of the third kind on I and I' which correspond to the image of ω in X_0.

Proposition 8.4. (*cf.* [S7]) *Let F be a cusp form of weight 2 for $\Gamma_1(Np)$ with coefficients in a finite extension K_F of $K = \mathbb{Q}_p(\zeta)$. Let $\mathcal{O}_F$ denote the ring of integers of K_F, and let $\omega_F = F(q)dq/q$ be the corresponding holomorphic differential on $X_1(Np)$ over K_F. Then ω_F is a regular differential on X over $\mathcal{O}_F$ if and only if the Fourier expansions*

$$(8.5) \qquad \begin{aligned} F(q) &= \sum_{n \geq 1} A_n q^n \quad \textit{and} \\ F|w_\zeta(q) &= \sum_{n \geq 1} B_n q^n \end{aligned}$$

both lie in $\mathcal{O}_F[[q]]$. In this case, let π_F be a uniformizing parameter in $\mathcal{O}_F$ and let a_n and b_n be the images of A_n and B_n in the residue field $\mathcal{O}_F/\pi_F\mathcal{O}_F$. Then $\omega_F \equiv (v, v')$ (mod $\pi_F L_F$) where

$$(8.6) \qquad \begin{aligned} v(q) &= \sum_{n \geq 1} a_n q^n dq/q \quad \textit{and} \\ v'|w_\zeta(q) &= \sum b_n q^n dq/q \end{aligned}$$

in a neighborhood of the cusp ∞ of $I = I_1(N)$.

Proof. Let ω be a regular differential on X. Since ∞ is a smooth section of X and q is a local parameter there, the pull-back of ω to Spec $\mathcal{O}[[q]]$ is a regular differential, so the expansion $\omega(q) = \Sigma A_n q^n dq/q$ must be integral. Since the automorphism w_ζ of X preserves L, the expansion of $\omega|w_\zeta(q) = \Sigma B_n q^n dq/q$ along ∞ must also be integral.

Conversely, assume ω_F is a holomorphic differential on $X_1(Np)$, and that the expansions $F(q)$ and $F|w_\zeta(q)$ are both integral. Then ω_F is a meromorphic section of $\Omega_{X/\mathcal{O}}$, which is an invertible sheaf on X. Hence the points where ω_F is not regular form a divisor D contained in the special fibre X_0. The integrality of the A_n shows that D does not intersect the section ∞, so D does not contain the component I. The integrality of the B_n shows that D does not intersect the section $w_\zeta(\infty)$, so D does not contain I'. Hence ω_F is a regular section of $\Omega_{X/\mathcal{O}}$.

The reduction of ω_F has the Fourier expansions $v = \Sigma\, a_n q^n dq/q$, $v'|w_\zeta = \Sigma\, b_n q^n dq/q$ at ∞, as the Fourier expansion of the reduction of a differential at ∞ is the reduction of the original Fourier expansion.

Proposition 8.7. *The endomorphisms $\langle d\rangle_p$, $\langle d\rangle_N$, T_l, U_l, U_p, and U_p' of the space of cusp forms of weight 2 for $\Gamma_1(Np)$ over K preserve the $\mathcal{O}$-lattice of regular differentials on X.*

Proof. This is clear for $\langle d\rangle_p$ and $\langle d\rangle_N$, which give automorphisms of X over $\mathcal{O}$. Let $t = T_l$ or U_l with $l \neq p$, and assume $\omega_F \in L$, so by Proposition 8.4 the Fourier coefficients of F and $F|w_\zeta$ lie in $\mathcal{O}$. By formulae (3.4), (3.5) and (3.7) the form $F|t$ has integral Fourier coefficients. By Proposition 6.7 $F|tw_\zeta = F|\langle l\rangle_p w_\zeta t$; the same formulae now show that $F|tw_\zeta$ has integral Fourier coefficients. Hence $\omega_F|t$ lies in L.

The Fourier expansions of $F|U_p$ and $F|U_p'$ have integral coefficients, by (3.7) and Proposition 6.10. (We recall that $w_{\zeta^d} = w_\zeta \cdot \langle d\rangle_p$.) Since $F|U_p w_\zeta = F|w_\zeta U_p'$ and $F|U_p' w_\zeta = F|w_\zeta U_p$, the same argument shows that the expansions of $F|U_p w_\zeta$ and $F|U_p' w_\zeta$ are integral. Hence $\omega_F|U_p$ and $\omega_F|U_p'$ both lie in L.

Let $\chi\colon (\mathbb{Z}/p\mathbb{Z})^\times \to \mathbb{Z}_p^\times$ be the Teichmüller character. The lattice L decomposes as a direct sum of $\mathcal{O}$-modues

$$L = \bigoplus_{j=1}^{p-1} L(j) \tag{8.8}$$

where

$$L(j) = \{\omega \in L\colon \omega|\langle d\rangle_p = \chi^j(d)\cdot\omega \text{ for all } d \in (\mathbb{Z}/p\mathbb{Z})^\times\}. \tag{8.9}$$

Since the automorphisms $\langle d\rangle_p$ commute with $\langle d\rangle_N$, T_l, U_l, U_p, and U_p', the latter endomorphisms of L preserve each eigenspace $L(j)$.

Proposition 8.10. 1) *$U_pU_p' = U_p'U_p = p\langle p\rangle_N$ on $\bigoplus_{j\neq p-1} L(j)$.*

2) *If $\omega \in \bigoplus_{j\neq p-1} L(j)$ then $\omega \equiv (v, v')$ (mod πL) where v and v' are holomorphic differentials on I and I', respectively.*

Proof. 1) This follows from Proposition 6.12, as the submodule $\bigoplus_{j\neq p-1} L(j)$ consists of the regular differentials ω_F such that $\Sigma_{d\in(\mathbb{Z}/p\mathbb{Z})^\times} F|\langle d\rangle_p = 0$.

2) Since v and v' are meromorphic, with poles of order $\leqslant 1$ at each supersingular point x, and $Res_x(v) + Res_x(v') = 0$ it suffices to prove that $Res_x(v) = 0$ for all $x \in \Sigma$.

But the group $\{\langle d\rangle_p : d \in (\mathbb{Z}/p\mathbb{Z})^\times\}$ fixes each supersingular point x on $I_1(N)$, as the Galois covering $I_1(N) \to X_1(N)$ over $\mathbb{Z}/p\mathbb{Z}$ is totally ramified at supersingular points. Hence $Res_x(v|\langle d\rangle_p) = Res_x(v)$ for all d. Since $\omega = \omega_F$ lies in $\bigoplus_{j\neq p-1} L(j)$, we have $\Sigma_d F|\langle d\rangle_p = 0$. Hence

$$0 = \sum_d Res_x(v|\langle d\rangle_p) = (p-1)Res_x v = -Res_x v$$

and v is regular at each supersingular point $x \in \Sigma$.

We now consider the corresponding eigenspaces of $\bar{L} = L/\pi L$: $\bar{L} = \bigoplus_{j=1}^{p-1} \bar{L}(j)$ with

$$(8.11) \qquad \bar{L}(j) = \{\omega \in \bar{L} : \omega|\langle d\rangle_p = d^j\omega\} = L(j)/\pi L(j).$$

Each $\bar{L}(j)$ is a finite dimensional vector space over $\mathcal{O}/\pi\mathcal{O} \xrightarrow{\sim} \mathbb{Z}/p\mathbb{Z}$, with a commuting family of endomorphisms given by the Hecke operators $\langle d\rangle_N$, T_l, U_l, and U_p. We wish to identify the Hecke module $\bar{L}(j)$ in terms of cusp forms for $\Gamma_1(N)$ over $\mathbb{Z}/p\mathbb{Z}$ of weights $k = j + 2$ and $p + 3 - k = p + 1 - j$.

First assume that $j \neq (p-1)$, so $k = j + 2$ satisfies $3 \leqslant k \leqslant p$. Any $\omega \in L(j)$ satisfies $\omega \equiv (v, v')$ mod $\pi L(j)$, where v and v' are holomorphic differentials on I and I' by part 2) of Proposition 8.10. We have $v \in H^0(I_1(N), \Omega^1_{I_1(N)})(k-2)$ and $v'|w_\zeta \in H^0(I_1(N), \Omega^1_{I_1(N)})(2-k)$. By Proposition 5.7, there are cusp forms $f \in M_k^0$ and $g \in M_{p+3-k}^0$ such that

$$(8.11) \qquad \begin{aligned} v &= f(q)dq/q \\ v'|w_\zeta &= g(q)dq/q \end{aligned}$$

in a neighborhood of the cusp ∞ on $I_1(N)$. We define a map of $\mathbb{Z}/p\mathbb{Z}$-vector spaces (for $3 \leqslant k \leqslant p$)

$$(8.12) \qquad \begin{aligned} \rho_k \colon \bar{L}(k-2) &\to M_k^0 \oplus M_{p+3-k}^0 \\ \omega &\mapsto (f, g). \end{aligned}$$

This is clearly an isomorphism, as ω (mod $\pi L(k-2)$) is completely determined by (v, v'), which may be arbitrary holomorphic differentials in the $(k-2)$ eigenspace (by our description of $\bar{L}$, the regular differentials on X_0). If $\omega = \omega_F$ with F as in Proposition 8.4, the Fourier expansion of f and g at ∞ are given by $f(q) = \Sigma\, a_n q^n$ and $g(q) = \Sigma\, b_n q^n$.

We define a new action of the Hecke operators on M_{p+3-k}^0 by having $\langle d\rangle_N$ act as usual, and having T_l, U_l, and U_p acts as $l^{k-2}T_l$, $l^{k-2}U_l$, and $p^{k-2}U_p = 0$, respectively. Denote this twisted Hecke module by $M_{p+3-k}^0[k-2]$.

PROPOSITION 8.13. (*cf.* [S7]) *Assume that* $3 \leqslant k \leqslant p$. *Then the map* ρ_k *defined in* (8.11–8.12) *gives an isomorphism of Hecke modules*:

$$\rho_k\colon \bar{L}(k-2) \xrightarrow{\sim} M_k^0 \oplus M_{p+3-k}^0[k-2].$$

Proof. We have observed that ρ_k is an isomorphism of vector spaces. Let $\omega = \omega_F$ and suppose $\rho_k(\omega) = (f, g)$. We must prove that

$$(8.14) \qquad \begin{cases} \rho_k(\omega_{F|\langle d\rangle_N}) = (f|\langle d\rangle_N, g|\langle d\rangle_N) \\ \rho_k(\omega_{F|T_l}) = (f|T_l, l^{k-2}g|T_l) \\ \rho_k(\omega_{F|U_l}) = (f|U_l, l^{k-2}g|U_l) \\ \rho_k(\omega_{F|U_p}) = (f|U_p, 0). \end{cases}$$

These all follow from a calculation of the Fourier expansions and the fact that $\chi^{k-2}(l) \equiv l^{k-2} (\text{mod } p)$. We will only prove the final formula, and leave the other calculations to the reader.

The reduction of $F|U_p(q)$ is clearly $f|U_p(q)$, so the first component of $\rho_k(\omega_{F|U_p})$ is correct. We must prove that the Fourier expansion of $F|U_p w_\zeta$ lies in $\pi\mathcal{O}[[q]]$, so the second component of $\rho_k(\omega_{F|U_p})$ is equal to zero. But $F|U_p w_\zeta = F|w_\zeta U_p' = G|U_p'$ where $G(q) \in \mathcal{O}[[q]]$. We now appeal to the formula in Proposition 6.10, which shows that

$$\begin{aligned} G|U_p'(q) &\equiv \sum_d G|w_{\zeta^d}(q\zeta^d) \quad (\text{mod } p) \\ &\equiv \sum_d G|w_{\zeta^d}(q) \qquad (\text{mod } \pi) \end{aligned}$$

as $\zeta^d \equiv 1 \ (\text{mod } \pi\mathcal{O})$ for all d. But $\Sigma_d G|w_{\zeta^d} = \Sigma\, G|\langle d^{-1}\rangle_p|w_\zeta = 0$ as G lies in an eigenspace for the group $\{\langle d\rangle_p\colon d \in (\mathbb{Z}/p\mathbb{Z})^\times\}$ with nontrivial character. Hence $F|U_p w_\zeta(q) = G|U_p'(q) \equiv 0 \ (\text{mod } \pi\mathcal{O}[[q]])$. A similar argument shows that:

$$(8.15) \qquad \rho_k(\omega_{F|U_p'}) = (0, g|U_p).$$

We observe that the map ρ_k depends on the choice of primitive p^{th} root of unity ζ used to identify the general fibre of X with $X_1(Np)$ (and so to identify L with an $\mathcal{O}$-lattice in the space of cusp forms of weight 2 for $\Gamma_1(Np)$ over F). If $\rho_k(\omega_F) = (f, g)$ and ρ_k' is the isomorphism associated to $\zeta' = \zeta^d$, we have $\rho_k'(\omega_F) = (f, d^{2-k}g)$. This follows from the identity $w_{\zeta^d} = \langle d\rangle_p^{-1} w_\zeta = w_\zeta \langle d\rangle_p$ proved in (6.7). In particular, the surjection of Hecke modules

$$(8.16) \qquad \begin{aligned} \rho\colon \bar{L}(k-2) &\twoheadrightarrow M_k^0 \\ \omega_F &\mapsto f \end{aligned}$$

is independent of the choice of ζ. The kernel of ρ may be canonically identified with the Hecke module $M^0_{p+3-k} \otimes \mu_p^{\otimes k-2}$, where T_l and U_l act on μ_p by $\langle l \rangle_p = l$ and U_p acts as zero, via the map $(g \otimes \zeta^{\otimes k-2}) \mapsto (0, v_g | w_\zeta)$.

We now consider the case when $j = p - 1$, so $k = p + 1$ and $p + 3 - k = 2$. If $\omega \subset \bar{L}(p-1)$ we have $\omega \equiv (v, v')$, where v is a meromorphic differential on $X_1(N)$ with poles of order $\leqslant 1$ at the supersingular points. By Proposition 5.10, there is a unique cusp form $f \in M^0_{p+1}$ such that $v = f(q)dq/q$ in a neighborhood of the cusp ∞. We define the map

$$\rho\colon \bar{L}(p-1) \twoheadrightarrow M^0_{p+1} \qquad (8.17)$$
$$\omega_F \mapsto f.$$

The kernel of ρ consists of the differentials $\omega = (0, v')$ with v' holomorphic on $X_1(N)$. Hence $v'|w_\zeta = g(q)dq/q$ for a unique cusp form $g \in M^0_2$; in this case, g is independent of the choice of ζ. We define the map $\mu\colon M^0_2 \to \bar{L}(p-1)$ by $\mu(g) = (0, v')$.

Proposition 8.18. *There is an exact sequence of Hecke modules*

$$0 \to M^0_2[p-1] \xrightarrow{\mu} \bar{L}(p-1) \xrightarrow{\rho} M^0_{p+1} \to 0.$$

Proof. The maps μ and ρ commute with the usual actions of $\langle d \rangle_N$, T_l, and U_l, by a computation of Fourier expansions at ∞. But on $\bar{L}(p-1)$ we have the formula

$$(v, v')|U_p = (v|U_p, -v|w)$$

where $w = w_{\zeta^d}$ for any $d \in (\mathbb{Z}/p\mathbb{Z})^\times$. This follows from the Fourier expansion of U'_p in Proposition 6.10 and the fact that $w_{\zeta^d} = w_\zeta = w$ on $L(p-1)$. Hence U_p annihilates the image of μ; since it also acts trivially on $M^0_2[p-1]$, this shows μ is a homomorphism of Hecke modules.

We note that, unlike 8.13, there are cases when the sequence of Hecke modules in Proposition 8.18 does not split. The case when $N = 11$ and $p = 3$ gives a nonsplit example.

§9. Lifting eigenforms to weight 2 in characteristic zero.

Let $f = \Sigma a_n q^n$ be a normalized cuspidal eigenform, which is a newform of type (k, ε) for $\Gamma_1(N)$ (mod p). Since the Hecke module M^0_k has finite dimension over $\mathbb{Z}/p\mathbb{Z}$, the Fourier coefficients a_n of f and the values of the character ε generate a finite extension field E of $\mathbb{Z}/p\mathbb{Z}$.

Let R be the integral closure of $\mathbb{Z}_p$ in an algebraic closure of $\mathbb{Q}_p$; let m_R be the maximal ideal of R and *fix* an embedding of E into the residue field R/m_R (which is an algebraic closure of $\mathbb{Z}/p\mathbb{Z}$). We say a newform $F = \Sigma A_n q^n$ on $\Gamma_1(N)$ or $\Gamma_1(Np)$ over R is a lifting of f if the Fourier coefficients A_n of F satisfy the congruence

$$A_n \equiv a_n \pmod{m_R} \qquad (9.1)$$

for all $n \geqslant 1$. We also insist that the character $\varepsilon_F = \varepsilon_N \cdot \varepsilon_p$ associated to F satisfies the congruence

$$\varepsilon_N(d) \equiv \varepsilon(d) \pmod{m_R} \tag{9.2}$$

for all $d \in (\mathbb{Z}/N\mathbb{Z})^\times$. The next Proposition shows that if f has weight k and $2 \leqslant k \leqslant p + 1$, it has a lifting to a newform F of weight 2 over R. In this way, many questions on modular forms (mod p) can be reduced to the study of forms of weight 2 in characteristic zero (as was noted by Shimura [Sh3]).

PROPOSITION 9.3. 1) *Let f be a newform of weight 2 for $\Gamma_1(N)$ over E. Then there is a lifting of f to a newform F of weight 2 for $\Gamma_1(N)$ over R.*

2) *Let f be a newform of weight k for $\Gamma_1(N)$ over E, with $3 \leqslant k \leqslant p + 1$. If $k = p + 1$ we assume further that the expansion $f(q)$ has filtration $p + 1$. Then there is a lifting of f to a newform F of weight 2 for $\Gamma_1(Np)$ over R with $\varepsilon_p = \chi^{k-2}$, where $\chi\colon (\mathbb{Z}/p\mathbb{Z})^\times \hookrightarrow \mathbb{Z}_p^\times$ is the Teichmüller character.*

Before giving the proof of Proposition 9.3, we make a few remarks. Since $\chi(d) \equiv d \pmod{p}$ we have the congruence

$$\varepsilon_p(d) \equiv d^{k-2} \pmod{m_R} \tag{9.4}$$

for all $d \in (\mathbb{Z}/p\mathbb{Z})^\times$. When $p \geqslant 5$, Serre [S9] has shown that one can first specify a lifting $\varepsilon_N\colon (\mathbb{Z}/N\mathbb{Z})^\times \to R^\times$ of the character ε, and then find a lifting F of f of type $(2, \varepsilon_N\chi^{k-2})$. In particular, one can show there is a lifting F of f with character ε_F of order prime to p. This statement is false for $p = 2$ and 3: the space of cusp forms of weight 2 for $\Gamma_1(13)$ has dimension 2, and the characters of both eigenforms have order 6 in characteristic zero (so have order 3 (mod 2) and 2 (mod 3)). For more details on this example, see §17. This is the reason why Serre modified his conjectures in [S8] to use Katz's definition of modular forms (mod p).

We emphasize that there is no unicity of the lifting F, even if the character $\varepsilon_F = \varepsilon_N \cdot \varepsilon_p$ is fixed. For example, the space of cusp forms of weight 2 for $\Gamma_1(23)$ with $\varepsilon = 1$ has dimension 2; the two eigenforms F and F' in characteristic zero are conjugate over $\mathbb{Q}(\sqrt{5})$ and have the same reduction f (mod 5).

Proof. 1) We consider the curve $X_1(N)$ over $\mathbb{Z}_p$, where it is smooth. Let L be the free $\mathbb{Z}_p$-module of cusp forms of weight 2 for $\Gamma_1(N)$ over $\mathbb{Z}_p$; by (2.4) we have an isomorphism $L \xrightarrow{\sim} H^0(X_1(N), \Omega^1_{X_1(N)/\mathbb{Z}_p})$. In particular, the quotient L/pL is canonically identified with M_2^0, the space of cusp forms of weight 2 for $\Gamma_1(N)$ over $\mathbb{Z}/p\mathbb{Z}$.

Let H be the commutative $\mathbb{Z}_p$-algebra of endomorphisms of L which is generated by the automorphisms $\langle d \rangle_N$ and the correspondences T_l and U_l of $X_1(N)$ (including T_p, as $p \nmid N$). The H/pH acts on L/pL, and by hypothesis has the eigenvector $v_f = f(q)dq/q$ on $L/pL \otimes E$. Let m be the maximal ideal of H which annihilates v_f; then m contains pH and $H/m = E$.

Since H is a free $\mathbb{Z}_p$-algebra of finite rank, we may choose a minimal prime ideal $\mathfrak{p}$ of H which is contained in m and satisfies $\mathfrak{p} \cap \mathbb{Z}_p = 0$. The quotient $H/\mathfrak{p}$ is then a local integral domain (of finite rank over $\mathbb{Z}_p$) with maximal ideal $m/\mathfrak{p}$ and residue field E. Choose an embedding $i\colon H/\mathfrak{p} \hookrightarrow R$ which gives the fixed embedding $E \hookrightarrow R/m_R$ of residue fields, and define $\varepsilon_N(d) = i(\langle d \rangle_N)$ in $R^\times$, $A_l = i(T_l)$ in R for $l \nmid N$, and $A_l = i(U_l)$ in R for $l|N$. By the definition of m, we have $\varepsilon_N(d) \equiv \varepsilon(d)$, and $A_l \equiv a_l \pmod{m_R}$.

The ideal $\mathfrak{p} \otimes \mathbb{Q}_p$ of $H \otimes \mathbb{Q}_p$ is in the support of the module $L \otimes \mathbb{Q}_p$ (on which $H \otimes \mathbb{Q}_p$ has a faithful representation). Hence there is a nonzero eigenvector ω_F in $L \otimes \bar{\mathbb{Q}}_p$ annihilated by $\mathfrak{p}$. Here F is cusp form of weight 2 for $\Gamma_1(N)$ over $\bar{\mathbb{Q}}_p$, which satisfies $F|\langle d \rangle_N = \varepsilon_N(d)$, $F|T_l = A_l \cdot F$ and $F|U_l = A_l \cdot F$. By (3.7) and (3.8) the first Fourier coefficient A_1 of F must be nonzero (or else the entire Fourier expansion of F is zero, hence $F = 0$). If we normalize our eigenvector by the condition that $A_1 = 1$, all of its Fourier coefficients lie in R (and are given by (3.9)). In particular, F is a lifting of f. Since f is a newform of level N, the same holds for F.

2) We consider the curve $X_1(Np)$ over $\mathbb{Q}_p(\zeta)$, where ζ is a primitive p^{th} root of unity, and let X be the regular model over $\mathbb{Z}_p[\zeta]$ discussed in §7. Let L be the $\mathbb{Z}_p[\zeta]$-module of regular differentials on X, and let H be the commutative $\mathbb{Z}_p[\zeta]$-algebra of endomorphisms of L which is generated by the automorphisms $\langle d \rangle_N$, $\langle d \rangle_p$ and the correspondences T_l, U_l, and U_p of $X_1(Np)$. (These endomorphisms of the space of cusp forms of weight 2 over $\mathbb{Q}_p(\zeta)$ preserve the lattice L by Proposition 8.7.) Then the algebra $H/(1-\zeta)H$ acts as endomorphisms of $\bar{L} = L/(1-\zeta)L$, the space of regular differentials on the special fibre X_0.

Assume first that $3 \leqslant k \leqslant p$. By Proposition 8.13 the eigenform f in M_k^0 gives an eigenvector $(v_f, 0)$ for $H/(1-\zeta)H$ acting on the space $\bar{L}(k-2) \otimes E$. Let m be the maximal ideal of H which annihilates this eigenvector, and choose a minimal prime $\mathfrak{p}$ of H contained in m and satisfying $\mathfrak{p} \cap \mathbb{Z}_p[\zeta] = 0$. Choose an embedding $i\colon H/\mathfrak{p} \hookrightarrow R$ which gives the fixed embedding $E \hookrightarrow R/m_R$ of residue fields, and define $\varepsilon_N(d) = i(\langle d \rangle_N)$, $\varepsilon_p(d) = i(\langle d \rangle_p)$, $A_l = i(T_l)$ for $l \nmid N$, $A_l = i(U_l)$ for $l|N$, and $A_p = i(U_p)$. Since the eigenvector $(v_f, 0)$ occurs in the subspace $\bar{L}(k-2)$ we have $\varepsilon_p(d) \equiv d^{k-2} \pmod{m_R}$, so $\varepsilon_p = \chi^{k-2}$. By the definition of m we have $\varepsilon_N(d) \equiv \varepsilon(d)$, $A_l \equiv a_l$, $A_p \equiv a_p \pmod{m_R}$.

The existence of a cusp form F over $\bar{\mathbb{Q}}_p$ with these eigenvalues for $H \otimes \mathbb{Q}_p$ then follows as in 1). (F is an eigenvector annihilated by $\mathfrak{p}$ in $L \otimes \bar{\mathbb{Q}}_p$.) If we normalize F by the condition $A_1 = 1$, it gives a newform of weight 2 on $\Gamma_1(Np)$ over R which reduces to $f \pmod{m_R}$. The level of F is divisible by N as f was assumed to be a newform, and is divisible by p as $\varepsilon_p \neq 1$.

Finally, assume that f has weight $k = p+1$ and $f(q)$ has filtration $p+1$. Define H and L as above; by Proposition 8.18 there is an eigenvector (v_f, v') in $\bar{L}(p-1)$ which maps to the eigenvector f for $H/(1-\zeta)H$ in M_{p+1}^0. Indeed, since the filtration of $f(q)$ is $p+1$, the eigenvalues of T_l and U_l on f do *not* occur in the submodule $M_2^0[p-1]$. The construction of a lifting F, with $\varepsilon_p = 1$, proceeds as above; we note that F is new at p (or else $f(q)$ would have filtration 2).

The ambiguity in the lifting of f to F results from the choice of a minimal prime $\mathfrak{p}$ of H contained in m, as well as the choice of an embedding of $H/\mathfrak{p}$ into R extending the fixed embedding of residue fields.

For the rest of this §, we assume that the weight k of the newform $f = \Sigma\, a_n q^n$ satisfies $2 \leqslant k \leqslant p$ and that $a_p \neq 0$. Define the integer k' by

$$k + k' = p + 1, \tag{9.6}$$

so $1 \leqslant k' \leqslant p - 1$. By Proposition 4.10, the series

$$f'(q) = \theta^{k'} f(q) = \Sigma n^{k'} a_n q^n \tag{9.7}$$

has filtration $k' + p + 1$ in $\tilde{M}_{k'+2} = \tilde{M}_{k+2k'}$. (If $a_p = 0$, this series has filtration $k' + 2$ when $k \neq 2$, and filtration 2 when $k = 2$.) If $g' = \theta^{k'-1} f$, then g' has filtration pk' in $M_{pk'}$, and

$$f'(q) = \theta g'(q). \tag{9.8}$$

PROPOSITION 9.9. 1) *The differential* $v_f = f(q)dq/q$ *is holomorphic on* $I_1(N)$. *It has an expansion of the form*

$$v_f = \left(\alpha_x z^{p-k} + \sum_{\substack{n \geqslant p-1 \\ n \equiv p-k \,(\mathrm{mod}\, p-1)}} \alpha_n z^n\right) dz$$

at each supersingular point x, *where* z *is a local parameter at* x. *The constant* α_x *is nonzero for at least one* $x \in \Sigma$.

2) *The differential* $v_{f'} = f'(q)dq/q$ *is meromorphic and exact on* $I_1(N)$. *It is regular outside* Σ, *and has an expansion of the form*

$$v_{f'} = \left(\beta_x z^{-1-k'} + \sum_{\substack{n \geqslant 0 \\ n \equiv -1-k' \,(\mathrm{mod}\, p-1) \\ n \not\equiv -1 \,(\mathrm{mod}\, p)}} \beta_n z^n\right) dz$$

at each supersingular point x. *The constant* β_x *is nonzero for at least one* $x \in \Sigma$.

Proof. This follows from Proposition 5.8, given the filtrations of $f(q)$ and $f'(q)$, and identity (9.8) which shows that $v_{f'} = dg'$ is exact. We note that the product $\alpha_x \beta_x$ is independent of the choice of uniformizing parameter z at x.

Now let $F = \Sigma A_n q^n$ be a lifting of f to a newform of weight 2 and character $\varepsilon_F = \varepsilon_N \cdot \chi^{k-2}$ on $\Gamma_1(N)$ (when $k = 2$) or on $\Gamma_1(Np)$ (when $3 \leqslant k \leqslant p$). Since $a_p \neq 0$, the lifted coefficient A_p is a unit in R. We will describe a cusp form $F' = \Sigma A'_n q^n$ of weight 2 and character $\varepsilon_{F'} = \varepsilon_N \cdot \chi^{k'}$ on $\Gamma_1(Np)$, which is a normalized eigenvector for the operators T_l, U_l, and U_p and reduces (mod m_R) to $f'(q)$.

When $k = 2$, let u denote the unique unit root of the quadratic equation

$$x^2 - A_p x + p\varepsilon_N(p) = 0$$

in R, and define

$$F'(q) = \Sigma A_n q^n - u\Sigma A_n q^{np}. \tag{9.10}$$

This is the expansion of the "old form" $F(\tau) - uF(p\tau)$ on $\Gamma_1(Np)$. When $3 \leqslant k \leqslant p$ we define F' by

$$F|w_\zeta = c_\zeta \cdot F'. \tag{9.11}$$

In this case, F' is a newform by Proposition 6.14. When $k = 2$, F' is a normalized eigenform for T_l, U_l, and U_p with eigenvalue $A'_p = p\varepsilon_N(p)/u$ but is *not* an eigenvector for U'_p. Indeed, on the 2-dimensional space with basis $\langle F(\tau), F(p\tau)\rangle$ the operators T_l and U_l act via the scalars A_l, and

$$\left\{\begin{aligned} U_p &= \begin{pmatrix} A_p & 1 \\ -p\varepsilon_N(p) & 0 \end{pmatrix} \\ w &= \begin{pmatrix} 0 & p^{-1} \\ p\varepsilon_N(p) & 0 \end{pmatrix}. \end{aligned}\right. \tag{9.12}$$

Proposition 9.13. 1) *The Fourier coefficients A'_n of F' lie in R, and satisfy the congruence*

$$F'(q) \equiv f'(q) = \theta^{k'}f(q) \pmod{m_R}.$$

2) *The differential $F'(q)dq/q$ is formally exact: there is a series $G(q) \in R[[q]]$ such that $dG(q) = F'(q)dq/q$.*

3) *The differential $(1 - \zeta)^{k'}\omega_{F'}$ on $X_1(Np)$ is a regular differential on the model X over R.*

Proof. When $k = 2$ part 1) follows directly from formula (9.10). Indeed, $F'(q) \equiv \Sigma_{(n,p)=1} a_n q^n$, as $u \equiv a_p$. When $3 \leqslant k \leqslant p$, the coefficients of F' for $(n, p) = 1$ are given by (6.15): $A'_n = A_n/\chi(n)^{k-2} \equiv a_n n^{k'}$. Since $A'_p = p\varepsilon_N(p)/A_p \equiv 0$, we have $F'(q) \equiv \Sigma n^{k'} a_n q^n$ as claimed.

To prove 2), we must check that $A'_n \equiv 0 \pmod{nR}$ for all $n \geqslant 1$ (as we may then define $G(q) = \Sigma_{n \geqslant 1}(A'_n/n)q^n$). Since R is n-divisible when $(n, p) = 1$, we need only consider the case when n is divisible by p. Write $n = p^r \cdot m$ where m is prime to p. Since F' is an eigenvector for T_l, U_l and U_p we have $A'_n = A'_m \cdot (A'_p)^r$, so it suffices to check that $A'_p \equiv 0 \pmod{pR}$. When $k = 2$ this follows from (9.10), as $A'_p = A_p - u = p\varepsilon_N(p)/u$. When $k > 2$ this follows from (6.16), which shows that $A'_p = p\varepsilon_N(p)/A_p$.

To prove 3), we will use the criterion of Proposition 8.4. Clearly the expansion $(1 - \zeta)^{k'} \cdot F'(q)$ lies in $R[[q]]$ and reduces to zero (mod m_R). We must also verify that the expansion $(1 - \zeta)^{k'}F'|w_\zeta(q)$ is integral. When $k = 2$ we use (9.12), which shows that $F'|w = p\varepsilon_N(p)F(p\tau) - p^{-1}uF(\tau)$. If we multiply this series by $(1-\zeta)^{k'}$, which has the same valuation as p, it becomes integral. When $3 \leqslant k \leqslant p$ we have $F'|w = c_\zeta^{-1} \cdot F|w^2 = c_\zeta^{-1} \cdot \varepsilon_N(p)\varepsilon_p(-1) \cdot F$. But the constant c_ζ has the same p-adic valuation as the Gauss sum $\Sigma\chi(d)^{-k'}\zeta^d$, by formula (6.17). This sum has the same valuation as $(1-\zeta)^{k'}$, by Stickelberger's theorem [L, Ch. I]. Hence, the differential $(1-\zeta)^{k'}\omega_{F'}$ is regular, with nonzero reduction on I', in all cases.

§10. The cases when $N \leqslant 4$. In §2 we made the hypothesis that $N > 4$, so that the objects (E, α) being classified have no automorphisms. When $N \leqslant 4$ there are automorphisms of the following type (cf. [D3]). Any pair (E, α) has the automorphism -1 for $N = 1$ and $N = 2$. If E_3 is an elliptic curve with an automorphism ζ_3 of order 3 and $\alpha_3\colon \mu_3 \hookrightarrow \ker(1-\zeta_3)$ is an embedding, then (E_3, α_3) has an automorphism by ζ_3, as does (E_3) in level 1. If E_4 is an elliptic curve with an automorphism ζ_4 of order 4 and $\alpha_4\colon \mu_2 \hookrightarrow \ker(1-\zeta_4)$ is an embedding, then (E_4, α_4) has an automorphism by ζ_4, as does (E_4) in level 1. Finally, one of the cusps on $X_1(4)$ has an automorphism of order 2.

When $N \leqslant 4$ we define $X_1(N)$ over $\mathbb{Z}[1/N]$ as the coarse moduli scheme associated to the stack $\mathscr{M}_H[1/N]$ [DR, 234–243]. The line bundle $\underline{\omega}^{\otimes 2}$ can be defined on $X_1(4)$, $\underline{\omega}^{\otimes 3}$ can be defined on $X_1(3)$, $\underline{\omega}^{\otimes 4}$ can be defined on $X_1(2)$, and $\underline{\omega}^{\otimes 12}$ on $X_1(1)$ [K2, 1.10]. For weights k divisible by 2 (respectively 3, 4, and 12) we can define holomorphic modular forms for $\Gamma_1(N)$ as sections of $\underline{\omega}^{\otimes k}$. For other weights, we must use the definition at the beginning of §2. However, if we remove the points with extra automorphisms, we can define $\underline{\omega}$ on the open curves of level $N = 3$ and $N = 4$ and $\underline{\omega}^{\otimes 2}$ on the open curves of level $N = 2$ and $N = 1$. Modular forms give holomorphic sections of an appropriate power, with possibly a fractional order zero at the deleted points.

Proposition 2.5, which studies the base change of modular forms of weight k over $\mathbb{Z}[1/N]$-algebras R, remains true for $N \geqslant 2$. It only holds for level $N = 1$ when 2 and 3 are invertible in R; the q-expansion principal of Proposition 2.7 holds in all cases [K2, 1.8–1.9].

The results of §3 go through for $N \leqslant 4$ without change. In §4 we used the universal curve $\underline{E}$ over $X_1(N)$ to show that the derivation θ exists. A different construction of θ for level $N = 1$ is given in [S4], and the remaining results in this section were proved for level $N = 1$ in [Sw].

The results in §5 are all valid for $N = 3$ and $N = 4$; they must be modified for $N = 1$ and $N = 2$, as the Igusa covering $I_1(N)$ has degree $(p-1)/2$ for p odd and Galois group $(\mathbb{Z}/p\mathbb{Z})^*/\langle \pm 1 \rangle$. Over $X_1(N)^h$ the cover is ramified only at ordinary points with $\mathrm{Aut}(E, \alpha) \neq \langle \pm 1 \rangle$. The line bundle $\underline{\omega}^{\otimes 2}$ can always be defined on $I_1(N)$; it has a canonical section "a^2" which vanishes to order 1 at each supersingular point and whose $((p-1)/2)$-power is the Hasse invariant A. The ring $\tilde{M}$ is graded by even integers (mod $p-1$) and the isomorphism of (5.6) becomes

$$\underline{\omega}^{\otimes 2}\left(\frac{p-3}{2}\underline{ss}\right) = \Omega^1_{I_1(N)}(cusps).$$

Proposition 5.7 holds for even k with $2 \leqslant k \leqslant p - 1$. In Proposition 5.8 we find that $f(q)$ has filtration k if and only if $\mathrm{ord}_x(\omega_f) \geqslant (p - 1 - k)/2$ at all supersingular points x, with equality holding for at least one x. The case when $N = 1$ is treated briefly in [S4, 1.3] and in more detail in [S7].

The results of §6 remain valid for $N \leqslant 4$. In §7 all results remain valid provided $Np > 4$ (i.e., $p > 3$ when $N = 1$). When $N = 1$ and $N = 2$ the results in §8 and §9 hold for k even and $2 \leqslant k \leqslant p + 1$.

§11. Galois representations associated to eigenforms (mod p). Let $\bar{\mathbb{Q}}$ be an algebraic closure of $\mathbb{Q}$, and let $\mathscr{G} = \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$. Let l be a rational prime, λ a place of $\bar{\mathbb{Q}}$ dividing l, and $\mathscr{G}_\lambda$ the corresponding decomposition subgroup in $\mathscr{G}$. Let σ_λ be a Frobenius element in $\mathscr{G}_\lambda$, which satisfies $\sigma_\lambda(\alpha) \equiv \alpha^l \pmod{\lambda}$. Then σ_λ is well defined modulo the inertia subgroup of $\mathscr{G}_\lambda$.

If $\rho\colon \mathscr{G} \to GL_n(E)$ is a linear representation of $\mathscr{G}$, we say ρ is unramified at l if ρ is trivial on the inertia subgroup of $\mathscr{G}_\lambda$. (This condition is independent of λ dividing l, as the inertia subgroups of different factors are conjugate in $\mathscr{G}$.) In this case, the element $\rho(\sigma_\lambda)$ is well-defined in $GL_n(E)$, and its conjugacy class $\rho(\sigma_l)$ depends only on ρ and l.

Proposition 11.1. (*Deligne*) *Let $f = \Sigma a_n q^n$ be a normalized eigenform of type (k, ε) for $\Gamma_1(N)$, which is defined over the finite field E of characteristic p. Then there is a continuous, semi-simple Galois representation*

$$\rho = \rho_f\colon \mathscr{G} \to GL_2(E), \quad \textit{where } \mathscr{G} = \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$$

which is unramified for all primes $l \nmid Np$ and where

$$(11.2) \qquad Tr\ \rho(\sigma_l) = a_l \qquad \det \rho(\sigma_l) = \varepsilon(l) l^{k-1}$$

for all $l \nmid Np$.

Before giving the proof of this Proposition, we make two remarks. It follows from the Čebotarev density theorem that a semi-simple continuous representation of $\mathscr{G} = \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ is determined, up to isomorphism, by the characteristic polynomials of Frobenius elements on a set of primes l of density 1. Hence the representation ρ_f is determined up to isomorphism by (11.2). Also, since the cyclotomic character $\chi\colon \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \twoheadrightarrow \mathrm{Aut}(\mu_p) \simeq (\mathbb{Z}/p\mathbb{Z})^\times$ is unramified for $l \neq p$ and satisfies $\chi(\sigma_l) = l$, we have:

$$(11.3) \qquad \det \rho_f = \varepsilon \cdot \chi^{k-1}.$$

Here we have identified ε with a character of $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ via the surjective homomorphism $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \twoheadrightarrow \mathrm{Aut}(\mu_N) \simeq (\mathbb{Z}/N\mathbb{Z})^\times$.

Proof. When f is an Eisenstein series, the representation ρ_f is reducible, and its existence follows from the theory of cyclotomic fields. Hence we may assume that f is a cusp form, of weight $k \geqslant 1$.

If g has weight $k = 1$, we may find a normalized eigenform f of weight $k = p$ with $a_l(f) = a_l(g)$ for all $l \neq p$. Indeed, take f to be an eigenvector for U_p in the span of Ag and V_pg. The existence of the representation ρ_f implies the existence of $\rho_g(=\rho_f)$.

If g has weight $k \geqslant 2$, there is an eigenform f of weight k in the range $2 \leqslant k \leqslant p+1$ such that $\theta^i f(q) = g(q)$ for some $0 \leqslant i \leqslant p-1$. This is a result due to Tate when $N = 1$ (cf. [J, §7]); a proof for arbitrary N can be found in [AS, Thm. 3.4, 3.5]. The existence of the representation ρ_f implies the existence of ρ_g $(=\rho_f \otimes \chi^i)$. Consequently, to prove Proposition 11.1 it suffices to demonstrate the existence of ρ_f for f a cusp form of weight k, with $2 \leqslant k \leqslant p+1$. If $k = p+1$ we may assume that $f(q)$ has filtration $= p+1$, for if $f = A \cdot q$ then g has weight 2 and $\rho_f = \rho_g$. Also, it is no loss of generality to assume that f is a newform on $\Gamma_1(N)$.

Let $F = \Sigma A_n q^n$ be a lifting of the eigenform f to a newform of weight 2 on $\Gamma_1(N)$ (when $k = 2$) or on $\Gamma_1(Np)$ (when $3 \leqslant k \leqslant p+1$). This lifting is guaranteed by Proposition 9.3. Let $\varepsilon_F = \varepsilon_N \cdot \chi^{k-2}$ be the character of F. We recall that the Fourier coefficients A_n of F, as well as the values of the character ε_F, lie in a finite integral extension $\mathcal{O}_K$ of $\mathbb{Z}_p$, with residue field E and quotient field K of finite degree over $\mathbb{Q}_p$. We now recall a standard construction of a continuous Galois representation ρ_F: $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to GL_2(K)$, using the p-power division points in the Jacobian.

PROPOSITION 11.4. *Let $F = \Sigma A_n q^n$ be a newform of weight 2 and character ε_F for $\Gamma_1(M)$, which is defined over a finite extension K of $\mathbb{Q}_p$. Then there is a continuous Galois representation*

$$\rho_f\colon \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to GL_2(K)$$

which is unramified for all primes $l \nmid Mp$, and where

$$Tr\ \rho_F(\sigma_l) = A_l \qquad \det \rho_F(\sigma_l) = \varepsilon_F(l) \cdot l$$

for all $l \nmid Mp$.

Proof. Let J be the Jacobian of $X_1(M)$ over $\mathbb{Q}$; this abelian variety has good reduction at all primes $l \nmid M$. The Hecke correspondences T_l, for $l \nmid M$, induce endomorphisms of J over $\mathbb{Q}$ (viewing J as the Albanese variety of $X_1(M)$) and the congruence of Proposition 3.12 shows that $T_l = \mathrm{Ver}_l + \langle l \rangle_M \mathrm{Fr}_l$ in $\mathrm{End}_{\mathbb{Z}/l\mathbb{Z}}(J)$. We also have the identity $\mathrm{Ver}_l \cdot \mathrm{Fr}_l = \mathrm{Fr}_l \cdot \mathrm{Ver}_l = l$ in the endomorphism ring of J over $\mathbb{Z}/l\mathbb{Z}$.

Let H be the commutative subring of $\mathrm{End}_{\mathbb{Q}}(J)$ which is generated over $\mathbb{Z}$ by the operators T_l, for l prime to M and the automorphisms $\langle d \rangle_M$ for $d \in (\mathbb{Z}/M\mathbb{Z})^*$. The canonical polarization of J induces an anti-involution $t \mapsto t^*$ of $\mathrm{End}(J)$ (the Rosati involution). On the subring H, we have $t^* = w_M t w_M$, where $w_M = w_{\zeta_M}$ is the involution of $X_1(M)$ over $\mathbb{Q}(\zeta_M)$ defined as in (6.4), using a primitive M^{th} root of unity [MW, Ch. 2, §5]. If $\sigma = \sigma_a$ in $\mathrm{Gal}(\mathbb{Q}(\zeta_M)/\mathbb{Q}) \simeq (\mathbb{Z}/M\mathbb{Z})^*$, we have $w_M^\sigma = \langle a \rangle_M^{-1} w_M = w_M \cdot \langle a \rangle_M$ as in part 2) of Proposition 6.7.

Let $T_pJ = \varprojlim_n J[p^n](\bar{\mathbb{Q}})$ be the Tate module of J, and let $V = T_pJ \otimes_{\mathbb{Z}_p} K$, which is a module for the K-algebra $H \otimes K$. The newform F gives a character of $H \otimes K$; we let $W \subset V$ be the subspace on which $H \otimes K$ acts via the character associated to F. It is well known that W has dimension 2 over K: this follows from the fact that the F-eigencomponent in the space of holomorphic differentials has dimension 1 (as F is a newform) and the representation of $H \otimes K$ on V is (via a comparison theorem in étale cohomology) the direct sum of its action on differentials and the dual representation [D1].

The group $\mathscr{G} = \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ also acts K-linearly on V, and this representation is unramified outside Mp. The Galois action commutes with the action of $H \otimes K$, as the endomorphisms generating H are defined over $\mathbb{Q}$. Hence $\mathscr{G}$ acts on W, and this gives a continuous representation

$$r_F\colon \mathscr{G} \to \mathrm{Aut}_K(W) = GL_2(K)$$

unramified outside Mp. We claim that the characteristic polynomial of $r_F(\sigma_l)$, for $l \nmid Mp$, is equal to

$$(11.5) \qquad x^2 - A_l/\varepsilon_F(l)x + l/\varepsilon_F(l).$$

Indeed, on W the operator $T_l = A_l$ and $\langle l \rangle_M = \varepsilon_F(l)$. By the Eichler-Shimura congruence, $A_l = l/r_F(\sigma_l) + \varepsilon_F(l)r_F(\sigma_l)$ in $\mathrm{End}_K(W)$. Hence $r_F(\sigma_l)$ satisfies (11.5). To prove this is its characteristic polynomial, it suffices to show that

$$(11.6) \qquad \det(\sigma_l | W) = l/\varepsilon_F(l).$$

To prove (11.6), we consider the Weil pairing [W]:

$$(\ ,\ )\colon T_pJ \times T_pJ \to T_p\mathbb{G}_m = \mathbb{Z}_p(1).$$

This is strictly alternating, and satisfies $(a^\sigma, b^\sigma) = (a, b)^\sigma$ for all $\sigma \in \mathscr{G}$, $(ta, b) = (a, t^*b)$ for all $t \in \mathrm{End}(J)$. Define another alternating form on T_pJ by:

$$\langle a, b \rangle = (a, w_M b).$$

Then $\langle ta, b \rangle = \langle a, tb \rangle$ for all $t \in H$, so $\langle\ ,\ \rangle$ induces a nondegenerate alternating pairing

$$\langle\ ,\ \rangle\colon W \times W \to K$$

which can be used to compute the Galois representation $\det r_F$. If $a, b \in W$ and $\sigma = \sigma_l$, we have $\langle a, b\rangle^\sigma = (a, w_M b)^\sigma = (a^\sigma, w_M^\sigma b^\sigma) = (a^\sigma, w_M \cdot \langle l \rangle_M b^\sigma) = \langle a^\sigma, \varepsilon_F(l) b^\sigma \rangle =$

$\varepsilon_F(l)\langle a^\sigma, b^\sigma\rangle$. On the other hand, $\langle a, b\rangle^\sigma = (a, w_M b)^\sigma = l\cdot(a, w_M b) = l\cdot\langle a, b\rangle$, so

$$\langle a^\sigma, b^\sigma\rangle = \langle a, b\rangle\cdot l/\varepsilon_F(l)$$

and $\det(\sigma_l | W) = l/\varepsilon_F(l)$ as claimed in (11.6).

We now define the representation ρ_F by

$$\rho_F = r_F \otimes \varepsilon_F. \tag{11.7}$$

Since the characteristic polynomial of $r_F(\sigma_l)$ is given by (11.5), the characteristic polynomial of $\rho_F(\sigma_l)$ is equal to

$$x^2 - A_l x + l\varepsilon_F(l).$$

This completes the proof of Proposition 11.4.

We now apply Proposition 11.4 (with $M = N$ or Np) to the lifting F of our eigenform f. Since the representation ρ_F is continuous, and $\mathcal{G} = \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ is compact, the image of ρ_F stabilizes an $\mathcal{O}_K$-lattice of rank 2. Hence ρ_F can be reduced modulo the maximal ideal m_K of $\mathcal{O}_K$, and we define ρ_f as the semi-simplification of this reduction. Then ρ_f is independent of the lattice chosen, by the Brauer-Nesbitt theorem. Since

$$A_l = a_l \pmod{m_K}$$

$$\varepsilon_F(l)\cdot l = \varepsilon_N(l)\chi^{k-2}(l)\cdot l \equiv \varepsilon(l)l^{k-1} \pmod{m_K}$$

the characteristic polynomial of $\rho_f(\sigma_l)$ is as claimed in (11.2).

Let $f = \Sigma\, a_n q^n$ be a normalized eigenform of type (k, ε) for $\Gamma_1(N)$ with coefficients in the finite field E of characteristic p. Assume the weight k of f satisfies $2 \leqslant k \leqslant p+1$, and that when $k = p+1$, $f(q)$ has filtration $= p+1$. When $k = 2$, let $V = V_f$ be the subspace of $J_1(N)[p](\bar{\mathbb{Q}}) \otimes E$ on which T_l acts by multiplication by the scalar a_l, for all $l \nmid Np$, and $\langle d\rangle_N$ acts by multiplication by $\varepsilon(d)$. When $2 \leqslant k \leqslant p+1$, let $V = V_f$ be the subspace of $J_1(Np)[p](\bar{\mathbb{Q}}) \otimes E$ on which T_l acts as multiplication by a_l, for all $l \nmid Np$, $\langle d\rangle_N$ acts by multiplication by $\varepsilon(d)$, and $\langle d\rangle_p$ acts by multiplication by d^{k-2}. It follows from the proof of (11.2) that V is a non-trivial E-subspace of the p-torsion in the Jacobian; it affords a representation of the group $\mathcal{G} = \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$.

PROPOSITION 11.8. *Assume that the representation $\rho_f\colon \mathcal{G} \to GL_2(E)$ is irreducible. Then* $\dim V_f = 2r$ *is even, and the semi-simplification of the representation $V_f \otimes \varepsilon\chi^{k-2}$ is isomorphic to r copies of ρ_f.*

Proof. The Eichler-Shimura congruence shows that the Frobenius element σ_l satisfies the quadratic polynomial $x^2 - (a_l/\varepsilon(l)l^{k-2})x + l/\varepsilon(l)l^{k-2}$ in $\mathrm{End}_E(V)$. Hence σ_l satisfies $x^2 - a_l x + \varepsilon(l)l^{k-1}$ in the endomorphism ring of $V' = V \otimes \varepsilon\chi^{k-2}$. Let α_l

and β_l be the roots of this polynomial, and assume α occurs with multiplicity s and β occurs with multiplicity t in the semi-simplification of V'.

Let $V^* = \mathrm{Hom}(V, E)$ be the dual representation of $\mathscr{G}$. Then the eigenvalues of σ_l on the semi-simplification of $V'' = V^* \otimes \chi$ are α with multiplicity t and β with multiplicity s. Hence the characteristic polynomial of σ_l on $V' \oplus V''$ is equal to $(x^2 - a_l x + \varepsilon(l)l^{k-1})^{\dim V}$. Since this is true for all $l \nmid Np$, the semi-simplification of $V' \oplus V''$ is isomorphic to dim V copies of the representation ρ_f (where the characteristic polynomial of σ_l is equal to $x^2 - a_l x + \varepsilon(l)l^{k-1}$). Since ρ_f is irreducible, V' has semi-simplification isomorphic to $r = (\dim V)/2$ copies of ρ_f.

Note. In the next section we will show, following Mazur, that $r = 1$ in most cases.

§12. The local Galois representation at p (ordinary case). Let $f = \Sigma\, a_n q^n$ be a normalized eigenform of type (k, ε), which is a newform for $\Gamma_1(N)$ and is defined over the finite field E of characteristic p. Let $\rho_f \colon \mathscr{G} = \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to GL_2(E)$ be the associated semi-simple Galois representation, which was constructed in the proof of Proposition 11.1. In this section we will study the restriction of ρ_f to a decomposition group $\mathscr{G}_p$ at p, in the "ordinary case" when a_p (the eigenvalue of U_p acting on f) is nonzero. By Proposition 4.12, it is no loss of generality to assume that f has weight k, with $2 \leqslant k \leqslant p + 1$.

The character $\chi\colon \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \twoheadrightarrow \mathrm{Aut}(\mu_p) = (\mathbb{Z}/p\mathbb{Z})^\times$ restricts to a character of a decomposition group $\mathscr{G}_p \simeq \mathrm{Gal}(\bar{\mathbb{Q}}_p/\mathbb{Q}_p)$ at p, which has order $(p-1)$ on the inertia subgroup. For any $\alpha \neq 0$ in E, we let $\lambda(\alpha)$ denote the unramified character of $\mathrm{Gal}(\bar{\mathbb{Q}}_p/\mathbb{Q}_p)$ which maps a Frobenius element Fr_p (in the sense of Artin) to α.

PROPOSITION 12.1. *Let $f = \Sigma\, a_n q^n$ be a normalized eigenform of type (k, ε) for $\Gamma_1(N)$ over E, with $2 \leqslant k \leqslant p+1$ and $a_p \neq 0$. Let W be the 2-dimensional E-vector space underlying the representation ρ_f of $\mathscr{G} = \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$. There is an exact sequence of $\mathscr{G}_p \simeq \mathrm{Gal}(\bar{\mathbb{Q}}_p/\mathbb{Q}_p)$-modules*

$$(12.2) \qquad 0 \to L \to W \to L' \to 0 \quad \textit{with}\ \dim L = \dim L' = 1.$$

The group $\mathrm{Gal}(\bar{\mathbb{Q}}_p/\mathbb{Q}_p)$ acts on L by the character $\chi^{k-1}\cdot\lambda(\varepsilon(p)/a_p)$ and on L' by the unramified character $\lambda(a_p)$.

In terms of matrices, Proposition 12.1 states that there is a basis for W such that the subgroup $\mathscr{G}_p$ acts via upper triangular matrices:

$$(12.3) \qquad \begin{pmatrix} \chi^{k-1}\cdot\lambda(\varepsilon(p)/a_p) & * \\ 0 & \lambda(a_p) \end{pmatrix}.$$

We remark that the line L where $\mathscr{G}_p$ acts by the character $\chi^{k-1}\cdot\lambda(\varepsilon(p)a_p)$ is unique, except in the case when $k = p$, $a_p^2 = \varepsilon(p)$, and the representation W is unramified at p.

Proposition 12.1 was first proved by Deligne [D2] in a letter to Serre; the proof uses the étale cohomology of nontrivial sheaves on the modular curve and works for all $k \geqslant 2$. We will give a different proof for weights $2 \leqslant k \leqslant p$, using the realization of ρ_f in the p-torsion of the Jacobian of $X_1(Np)$. This proof is modeled on an argument due to Serre and Fontaine, who also treated the case when $a_p = 0$ (at least for $N = 1$) in an exchange of letters [F3, S7]. In our proof, we will assume f is a newform of level N and that the representation ρ_f is irreducible over E. When ρ_f is reducible, (12.1) follows from work of Swinnerton-Dyer [Sw].

A further question, posed by Serre, is whether or not the sequence of $\mathscr{G}_p$-modules in (12.2) is split. If f has filtration $= p + 1$, this sequence is always *nonsplit*: results of Mazur [*R*, §6] show that $a_p^2 = \varepsilon(p)$ and that ρ_f is "très ramifié" in the sense of Serre [S8, pg. 186]. When f has weight k, with $2 \leqslant k \leqslant p$, the splitting of (12.2) is a subtle question, intimately related to Serre's general conjectures on Galois representations and modular forms (mod p), which we will pursue in the next four sections.

Before beginning the proof of Proposition 12.1, we recall some basic results from the theory of p-divisible groups. (We refer to the papers of Tate [T] and Fontaine [F2] for the proofs.) Let $G = \varinjlim_n G_n$ be a p-divisible group of height h over the ring R. Let ${}^tG$ be its Cartier dual: $\varinjlim_n \operatorname{Hom}(G_n, \mathbb{G}_m)$, which is also p-divisible of height h over R.

If $R = K$ is a field of characteristic zero, with algebraic closure $\bar{K}$, the group G is étale and is completely determined by the Galois module $T_pG = \varprojlim G_n(\bar{K}) = \operatorname{Hom}_{\bar{K}}(\mathbb{Q}_p/\mathbb{Z}_p, G)$. This module is free of rank h over $\mathbb{Z}_p$, and the Galois group $\operatorname{Gal}(\bar{K}/K)$ acts continuously and $\mathbb{Z}_p$-linearly on T_pG. Conversely, any such Galois module determines a p-divisible group G over K. There is a canonical, nondegenerate $\mathbb{Z}_p$-linear pairing: $T_pG \times T_p{}^tG \to T_p\mathbb{G}_m = \mathbb{Z}_p(1)$.

If $R = L$ is a perfect field of characteristic p, the group G splits as the product $G^0 \times G^e$ of its connected and étale subgroups. The subgroup G^0 is also a product $G^m \times G^{ll}$ of its subgroups of multiplicative and local-local type: G^m is the largest subgroup of G whose dual ${}^tG^m$ is étale. We let $D(G)$ be the contravariant Dieudonné module of G, defined as in Fontaine [F1], [F2]. Then $D(G)$ is a free module of rank h over W_L, the Witt vectors of L; it has semi-linear endomorphisms F and V which satisfy $FV = VF = p$. A group G is étale if F acts invertibly on $D(G)$, and is multiplicative if V acts invertibly on $D(G)$. In general, we have $D(G^e) = \operatorname{Hom}_{\mathbb{Z}_p}(T_pG, W_L)$ and $D(G^m) = T_p{}^tG \otimes W_L$ [Br].

Finally, assume that R is a complete discrete valuation ring with quotient field K of characteristic zero and residue field L perfect of characteristic p. Let G be a p-divisible group over R and let $\bar{G}$ be the corresponding p-divisible group over L. The splitting $\bar{G} = \bar{G}^0 \times \bar{G}^e$ is reflected in an exact sequence:

$$0 \to G^0 \to G \to G^e \to 0$$

of p-divisible groups over R. This gives an exact sequence $0 \to T_pG^0 \to T_pG \to T_pG^e \to 0$ of $\operatorname{Gal}(\bar{K}/K)$-modules, where the Galois action on the quotient T_pG^e is unramified. If G_K is a p-divisible group over K, we say that G_K has good reduction over R if it is the general fibre of a p-divisible group G over R. Tate's fundamental theorem: $\mathrm{Hom}_R(G, G') = \mathrm{Hom}_K(G_K, G'_K)$ shows that G is uniquely determined by G_K, if it exists [T].

We begin the proof of Proposition 12.1 by defining a p-divisible group G over $\mathbb{Q}$ attached to a normalized eigenform f (mod p). When f has weight $k = 2$, we let J be the Jacobian of the curve $X_1(N)$ and let H be the commutative subring of $\mathrm{End}_{\mathbb{Q}}(J)$ which is generated over $\mathbb{Z}$ by the Hecke operators T_l for $l \nmid N$, U_l for $l|N$, and $\langle d\rangle_N$ for $d \in (\mathbb{Z}/N\mathbb{Z})^\times$. When f has weight k satisfying $3 \leqslant k \leqslant p$, we let J be the Jacobian of $X_1(Np)$ and H the commutative subring of $\mathrm{End}_{\mathbb{Q}}(J)$ generated by the operators T_l for $l \nmid N$, U_l for $l|N$, U_p, $\langle d\rangle_N$ for $d \in (\mathbb{Z}/N\mathbb{Z})^\times$, and $\langle d\rangle_p$ for $d \in (\mathbb{Z}/p\mathbb{Z})^\times$.

Proposition 12.4. *There is a maximal ideal $m = m_f$ of H with residue field $H/m = E$ such that the following congruences hold*:

$$T_l \equiv a_l$$

$$U_l \equiv a_l$$

$$\langle d\rangle_N \equiv \varepsilon(d) \qquad (\textit{modulo } m)$$

$$\textit{when } k \neq 2 \begin{cases} U_p \equiv a_p \\ \langle d\rangle_p \equiv d^{k-2} \end{cases} \quad (\textit{modulo } m).$$

Proof. Let F be a lifting of f to a newform of weight 2 on $\Gamma_1(N)$ or $\Gamma_1(Np)$, as guaranteed by Proposition 9.3. Then F gives a ring homomorphism $H \to R$, where R is the integral closure of $\mathbb{Z}_p$. The resulting homomorphism $H \to R \to R/m_R$ has kernel m, and is independent of the choice of lifting F.

Since H is free of finite rank over $\mathbb{Z}$, the ring $H_p = \varprojlim H/p^nH = H \otimes \mathbb{Z}_p$ is a complete semi-local $\mathbb{Z}_p$-algebra of finite rank. Similarly, the ring $H_m = \varprojlim H/m^nH$ is complete and local, with residue field $H_m/mH_m = H/m = E$. The maximal ideals of H_p correspond bijectively to the maximal ideals of the finite ring H/pH. In particular, m is a maximal ideal of H_p; by the theory of complete, semi-local rings H_m is a direct factor of H_p. This splitting gives an idempotent decomposition of the identity:

$$\text{(12.5)} \qquad \begin{aligned} H_p &= H_m \times H'_m \\ 1 &= \varepsilon_m + \varepsilon'_m \end{aligned}$$

which may be used to decompose any H_p-module (cf. [M, II §7]).

The structure of H_m is somewhat clarified by considering the set of all liftings F of f to R, the integral closure of $\mathbb{Z}_p$. (We recall that an identification of E with a subfield of R/m_R has been fixed.) Associated to the lifting F_i we have the order $R_i = \mathbb{Z}_p[A_n]$ generated by its Fourier coefficients, with residue field $R_i/m_i = E$, together with a surjective homomorphism $H_m \to R_i$ taking T_l to A_l, etc. The inertia group of $\mathbb{Q}_p$ acts on the set of liftings $\{F_i\}$ by conjugation of coefficients, and we have an injective homomorphism

$$(12.6) \qquad H_m \hookrightarrow \prod_{\text{orbits}} R_i$$

whose image is contained, with finite index, in the subring of elements $(r_1, \ldots, r_n)$ with $r_i \equiv r_j$ in E (the index reflects higher congruences among the orbits of liftings). In particular, the artinian $\mathbb{Q}_p$-algebra $H_m \otimes \mathbb{Q}_p$ is a product of fields, and the newform F lifting f in Proposition 9.3 is unique if and only if H_m is unramified over $\mathbb{Z}_p$. We define

$$(12.7) \qquad h = \operatorname{rank}_{\mathbb{Z}_p}(H_m) = \dim_{\mathbb{Q}_p}(H_m \otimes \mathbb{Q}_p).$$

We note that T_p is a unit in H_m, when $k = 2$, as $T_p \equiv a_p \neq 0$ (modulo m). Similarly, U_p is a unit in H_m when $3 \leqslant k \leqslant p$.

Let $T_pJ = \varprojlim_n J[p^n](\bar{\mathbb{Q}})$ be the Tate module of the Jacobian, which is a module for H_p. The $\mathbb{Z}_p$-module $\varepsilon_m T_pJ$ (where ε_m is the idempotent for H_m defined in (12.5)) is free over $\mathbb{Z}_p$, and is stable under the action of $\mathscr{G} = \operatorname{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$. It therefore defines a p-divisible group G over $\mathbb{Q}$ with $T_pG = \varepsilon_m T_pJ$. This is the p-divisible group associated to the eigenform f (mod p); we note that G has endomorphisms by H_m over $\mathbb{Q}$. The representation of $\mathscr{G}$ on $G[m^n](\bar{\mathbb{Q}})$ is given by the Galois action on T_pG/m^nT_pG. We will prove that T_pG is a free H_m-module of rank 2 (in most cases) and that the representation of $\mathscr{G}$ on T_pG/mT_pG is isomorphic to the representation $\rho_f \otimes (\varepsilon \cdot \chi^{k-2})^{-1}$. The structure of G over $\mathbb{Q}_p$ will allow us to determine the restriction of ρ_f to $\mathscr{G}_p$. We begin with the case when f has weight 2.

PROPOSITION 12.8. *Assume that the newform f has weight $k = 2$.*

1) *The p-divisible group G has height $2h$ and is isomorphic to ${}^tG$ over $\mathbb{Q}(\mu_N)$. It has good reduction over $\mathbb{Z}_p$.*

2) *Let $\bar{G}$ be the reduction of G over $\mathbb{Z}/p\mathbb{Z}$, and $D(\bar{G})$ its Dieudonné module. The endomorphism F of $D(\bar{G})$ is $\mathbb{Z}_p$-linear, commutes with H_m, and satisfies the quadratic polynomial $\langle p\rangle_N \cdot x^2 - T_px + p = 0$ in $\operatorname{End}(D(\bar{G}))$.*

3) *$\bar{G} \overset{\sim}{\to} \bar{G}^m \times \bar{G}^e$, where the multiplicative and étale components of $\bar{G}$ each have height h over $\mathbb{Z}/p\mathbb{Z}$. The endomorphism F of $D(\bar{G}^e)$ acts via multiplication by a unit u in H_m which satisfies $u \equiv a_p/\varepsilon(p)$ (modulo m). The endomorphism V of $D(\bar{G}^m)$ acts via multiplication by the unit $u \cdot \langle p\rangle_N$ in H_m.*

4) *There is an exact sequence of p-divisible groups $0 \to G^0 \to G \to G^e \to 0$ over $\mathbb{Z}_p$, where G^0 is of multiplicative type and both G^0 and G^e have height h. The Galois group $\mathscr{G}_p = \operatorname{Gal}(\bar{\mathbb{Q}}_p/\mathbb{Q}_p)$ acts on T_pG^0 by the character $\lambda(u^{-1} \cdot \langle p\rangle_N^{-1}) \cdot \chi_p$ (where χ_p is the p-adic cyclotomic character giving the action on $T_p\mathbb{G}_m$) and acts on T_pG^e by the unramified character $\lambda(u)$.*

5) *There is an exact sequence of E-vector space schemes*

$$0 \to G^0[m] \to G[m] \to G^e[m] \to 0$$

over $\mathbb{Z}_p$. The connected component $G^0[m]$ has dimension $d^0 \geqslant 1$ and the étale compo-

nent $G^e[m]$ has dimension $d^e = 1$. The Galois group $\mathscr{G}_p$ acts on the semi-simplification of $G^0[m]$ via the character $\lambda(1/a_p)\cdot\chi$ (with multiplicity d^0) and on $G^e[m]$ via the character $\lambda(a_p/\varepsilon(p))$.

Proof. 1) The height of G is equal to the dimension of the $\mathbb{Q}_p$-vector space $V_pG = T_pG \otimes \mathbb{Q}_p = \varepsilon_m \cdot V_pJ$. But V_pG is a free $H_m \otimes \mathbb{Q}_p$ module of rank 2, by the theorem of multiplicity 1 for GL_2, as every lifting F of f is a newform (cf. [M, II §6–7]). Hence it has $\mathbb{Q}_p$-dimension $2h$.

The alternating form $\langle\ ,\ \rangle$ defined in the proof of (11.6) introduces a nondegenerate pairing of $\mathbb{Z}_p$-modules $T_pG \times T_pG \to T_p\mathbb{G}_m$. If $\sigma = \sigma_l$ with $l \equiv 1 \pmod{N}$, we have $\langle a^\sigma, b^\sigma\rangle = l\langle a, b\rangle$ for all $a, b \in T_pG$. Since such elements are dense in $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}(\mu_N))$, G is isomorphic to ${}^tG$ over $\mathbb{Q}(\mu_N)$.

Since J is an abelian scheme over $\mathbb{Z}[1/N]$, the p-divisible group $\varinjlim J[p^n]$ has good reduction over $\mathbb{Z}_p$. The same holds for its subgroup G, which is the limit of finite group-schemes $\varinjlim J[m^n]$ over $\mathbb{Z}_p$.

2) By part 1), $D(\bar{G})$ is a free $\mathbb{Z}_p$-module of rank $2h$, with endomorphisms by H_m, and $D(\bar{G}) \otimes \mathbb{Q}_p$ is free of rank 2 over $H_m \otimes \mathbb{Q}_p$. The endomorphisms F and V are $\mathbb{Z}_p$-linear, as $\bar{G}$ is defined over $\mathbb{Z}/p\mathbb{Z}$ where p^{th}-power is the identity. The Eichler-Shimura congruence: $T_p = \langle p\rangle_N F + V$ shows that F satisfies $\langle p\rangle_N x^2 - T_p x + p = 0$ in $\mathrm{End}(D(\bar{G}))$.

3) Since $T_p \equiv a_p \neq 0$ (modulo m), T_p is a unit in H_m and the quadratic polynomial satisfies by F factors over H_m: $(x^2 - \langle p\rangle_N^{-1} T_p x + \langle p\rangle_N^{-1} p) = (x-u)(x-u')$, where u is a unit in H_m. We have $u \equiv a_p/\varepsilon(p)$ (modulo m) and $u' = p/u \cdot \langle p\rangle_N$. Hence the eigenvalues of F on $D(\bar{G})$ are either units or divisible by p, and $\bar{G}$ has no local-local part. Since $\bar{G}$ is self-dual over $\mathbb{Z}/p\mathbb{Z}[\mu_N]$ by 1), we must have $\mathrm{height}(\bar{G}^m) = \mathrm{height}(\bar{G}^e) = h$. Hence F acts on $D(\bar{G}^e)$ by multiplication by u, and V acts on $D(\bar{G}^m)$ by multiplication by $p/u' = u\langle p\rangle_N$.

4) The exact sequence of groups over $\mathbb{Z}_p$ follows immediately from 3), and the general theory of p-divisible groups. The characters of $\mathscr{G}_p$ on $T_p(G^0)$ and $T_p(G^e)$ follow from the eigenvalues of F on $D(\bar{G}^m)$ and $D(\bar{G}^e)$ respectively.

5) Everything is an immediate consequence of 4), except that we may only conclude that d^0 and d^e are $\geqslant 1$. (The H_m-modules $T_p(G^0)$ and $T_p(G^e)$ are nontrivial, as they become free $H_m \otimes \mathbb{Q}_p$ modules when tensored with $\mathbb{Q}_p$.) To show that $d^e = 1$, we follow the method of Mazur [M, pg. 118–119]. The E-vector space $\bar{G}^e(m)$ is a subgroup of the p-torsion of $J_1(N)$ over the algebraic closure of $\mathbb{Z}/p\mathbb{Z}$, on which the algebra H acts via the eigenvalues of the form f. We will bound this subgroup, by studying its image in the differentials of the first kind.

Recall that there is a natural injection δ from the p-torsion in $\mathrm{Pic}^0(X_1(N))$ to the space $H^0(X_1(N), \Omega^1_{X_1(N)})$ over an algebraic closure of $\mathbb{Z}/p\mathbb{Z}$: if D is a divisor with $pD = (f)$ we have $\delta(D) = df/f$. The map δ is a twisted homomorphism of Hecke modules [Wi, Prop. 6.5], namely $\delta(tD) = t^*\delta(D)$ for all $t \in H$, where $t^* = w_N t w_N$. (This twist arises from the fact that we have defined the action of the Hecke correspondences of $X_1(N)$ on the Jacobian $J = J_1(N)$ by considering J as the Albanese (*not* the Picard) variety of the curve.) Hence the image of $\delta(\bar{G}^e(m))$ lies in the space of holomorphic differentials where $w_N t w_N = t^*$ acts by the eigenvalue of t on ω_f, for all $t \in H$. This space is 1-dimensional over E, and spanned by $\omega_{f|w_N}$, by the q-expansion principle. Hence $d^e \leqslant 1$.

The analogous results when f has weight k with $3 \leqslant k \leqslant p$ are obtained using the geometry of $X_1(Np)$, instead of the curve $X_1(N)$.

PROPOSITION 12.9. *Assume that f has weight k with $3 \leqslant k \leqslant p$.*

1) *The p-divisible group G has height $2h$ and is isomorphic to ${}^tG$ over $\mathbb{Q}(\mu_{Np})$. It has good reduction over the extension $\mathbb{Z}_p[\zeta_p]$ of $\mathbb{Z}_p$.*

2) *Let $\bar{G}$ be the reduction of G over $\mathbb{Z}_p[\zeta_p]/(1-\zeta_p)\mathbb{Z}_p[\zeta_p] = \mathbb{Z}/p\mathbb{Z}$, and $D(\bar{G})$ its Dieudonné module. Then $\bar{G} = \bar{G}^m \times \bar{G}^e$, where the multiplicative and étale components of $\bar{G}$ each have height h over $\mathbb{Z}/p\mathbb{Z}$. The endomorphism F of $D(\bar{G})$ is $\mathbb{Z}_p$-linear and commutes with the action of H_m. F acts on $D(\bar{G}^e)$ via multiplication by the unit $U_p \cdot \langle p \rangle_N^{-1}$ of H_m, and V acts on $D(\bar{G}^m)$ via multiplication by the unit U_p of H_m.*

3) *The exact sequence $0 \to G^0 \to G \to G^e \to 0$ of p-divisible groups over $\mathbb{Z}_p[\zeta_p]$ gives a filtration $0 \to T_pG^0 \to T_pG \to T_pG^e \to 0$ of T_pG which is stable under the group $\mathscr{G}_p = \mathrm{Gal}(\bar{\mathbb{Q}}_p/\mathbb{Q}_p)$. The Galois group $\mathscr{G}_p$ acts on T_pG^0 by the character $\lambda(U_p^{-1}) \cdot \chi_p$, where χ_p is the p-adic cyclotomic character, and on T_pG^e by the character $\lambda(U_p \cdot \langle p \rangle_N^{-1}) \cdot \chi^{2-k}$.*

4) *There is an exact sequence of E-vector space schemes*

$$0 \to G^0[m] \to G[m] \to G^e[m] \to 0$$

over $\mathbb{Q}_p$, with flat extensions to $\mathbb{Z}_p[\zeta_p]$. The connected component $G^0[m]$ has dimension $d^0 \geqslant 1$ and the étale component has dimension $d^e = 1$. The Galois group $\mathscr{G}_p$ acts on the semi-simplification of $G^0[m]$ via the character $\lambda(1/a_p) \cdot \chi$ (with multiplicity d^0) and on $G^e[m]$ via the character $\lambda(a_p/\varepsilon(p)) \cdot \chi^{2-k}$.

Proof. 1) The height and duality statements are proved exactly as in Proposition 12.8. In this case, however, G does *not* have good reduction over $\mathbb{Z}_p$, but has potentially good reduction (over the ramified extension $\mathbb{Z}[\zeta_p]$ it achieves good reduction). Indeed, the Jacobian J of $X_1(Np)$ is isogenous to the product of abelian varieties $A \times B$, where A is the connected component of the subgroup of points P in J with $\Sigma_{d \in (\mathbb{Z}/p\mathbb{Z})^\times} \langle d \rangle_p P = 0$, and B is the connected component of the subgroup of points fixed by the group $\{\langle d \rangle_p : d \in (\mathbb{Z}/p\mathbb{Z})^\times\}$. The isogeny $\varphi: J \to A \times B$ defined by $\varphi(P) = ((p-1)P - \Sigma\langle d \rangle_p P, \Sigma\langle d \rangle_p P)$ has degree prime to p (so induces an isomorphism of p-divisible groups), as the composite of φ with the canonical inclusion is multiplication by $(p-1)$ on J. The group G is a subgroup of the p-divisible group of A, as $\langle d \rangle_p = \chi^{k-2}(d)$ in $\mathrm{End}(G)$. In fact, the Hecke algebra H stabilizes A, and $T_pG = \varepsilon_m T_pA$ where ε_m is the idempotent of H_m defined in (12.5). An important result of Deligne and Rapoport [DR, V, Thm. 3.2] shows that the abelian variety A (and hence the p-divisible group G) has good reduction over $\mathbb{Z}_p[\zeta_p]$.

2) Let I and I' be the components of the special fibre X_0 of the model X for $X_1(Np)$ over $\mathbb{Z}_p[\zeta_p]$ described in §7. A result of Raynaud [Ra, Thm. 12.1] gives an isomorphism from $\mathrm{Pic}^0(X/\mathbb{Z}_p[\zeta_p])$ to the connected component of the Néron model of $J = J_1(Np)$ over $\mathbb{Z}_p[\zeta_p]$ (which is isogenous to the product of the Néron models

of A and B). Standard arguments (cf. [MW, pg. 268–269]) then show that the reduction $\bar{G}$ is isomorphic to the p-divisible group $\varinjlim_n \mathrm{Jac}(I) \times \mathrm{Jac}(I')[m^n]$ over $\mathbb{Z}/p\mathbb{Z}$.

This group splits as a product of $\bar{G}^m = \varinjlim_n \mathrm{Jac}(I)[m^n]$ by $\bar{G}^e = \varinjlim_n \mathrm{Jac}(I')[m^n]$. Indeed, the m^n-torsion in $\mathrm{Jac}(I)$ is clearly a subgroup (as H acts on I); since $U_p \equiv a_p \neq 0$ is a unit in H_m, and $U_p = \mathrm{Ver}_p$ on I, this subgroup is of multiplicative type (as V is a unit on its Dieudonné module). The element U_p in H does not act on I', but it acts on the subgroup of $\mathrm{Jac}(I')$ where $\Sigma\langle d\rangle_p = 0$ by formula (7.4). Hence it makes sense to speak of the m^n-torsion in $\mathrm{Jac}(I')$; since U_p is a unit in H_m and acts on the Dieudonné module of this subgroup by $\langle p\rangle_N \cdot F$, it must be étale. Both subgroups have height h, as $\bar{G}$ is self dual over $\mathbb{Z}/p\mathbb{Z}[\mu_N]$ by 1).

3) The filtration $0 \to T_pG^0 \to T_pG \to T_pG^e \to 0$ is clearly stable under the normal subgroup $\mathrm{Gal}(\bar{\mathbb{Q}}_p/\mathbb{Q}_p(\mu_p))$ of $\mathcal{G}_p$, which acts via the characters $\lambda(U_p^{-1})\cdot\chi_p$ on T_pG^0 and $\lambda(U_p\cdot\langle p\rangle_N^{-1})$ on T_pG^e. Since these characters are nonconjugate (one is unramified, and the other an unramified twist of χ_p), the filtration is stable under the action of $\mathcal{G}_p$ on T_pG. To determine the characters of $\mathcal{G}_p$ on T_pG^0 and T_pG^e, we recall that the inertia group $\mathrm{Gal}(\mathbb{Q}_p(\mu_p)/\mathbb{Q}_p)$ acts trivially on I (and hence on $\bar{G}^m$) and by $\langle d\rangle_p^{-1}$ on I' (and hence by $(\chi^{k-2})^{-1}$ on $\bar{G}^e$): this is the content of Proposition 7.2. This gives the claim.

4) This follows from 3), which gives $d^0, d^e \geqslant 1$. To show $d^e = 1$ we argue exactly as in the proof of Proposition 2.8. We leave the details to the reader.

PROPOSITION 12.10. (*cf.* [M, II §15]) *Assume that the newform* $f = \Sigma a_n q^n$ *is an eigenform of weight* $2 \leqslant k \leqslant p$ *with* $a_p \neq 0$, *and that the representation* $\rho_f\colon \mathcal{G} = \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to GL_2(E)$ *is irreducible. When* $k = p$ *assume further that* $a_p^2 \neq \varepsilon(p)$. *Then*

1) *We have* $d^0 = \dim_E G^0[m] = 1$ *and the representation of* $\mathcal{G}$ *on* $G[m]$ *is isomorphic to* $\rho_f \otimes (\varepsilon\cdot\chi^{k-2})^{-1}$.

2) T_pG *is a free* H_m*-module of rank* 2.

3) H_m *is a Gorenstein ring.*

Proof. 1) By Proposition 12.8 and 12.9 we know that the characters of $\mathcal{G}_p$ on the semi-simplification of $G[m]$ are $\lambda(1/a_p)\cdot\chi$, with multiplicity d^0, and $\lambda(a_p/\varepsilon(p))\chi^{2-k}$, with multiplicity $d^e = 1$. Our hypothesis that $a_p^2 \neq \varepsilon(p)$ when $k = p$ implies that these two characters of $\mathcal{G}_p$ are distinct. But Proposition 11.8 implies that the semi-simplification of $G[m]$ as a $\mathcal{G}$-module is isomorphic to r copies of the representation $\rho_f \otimes (\varepsilon\cdot\chi^{k-2})^{-1}$. Hence each character of $\mathcal{G}_p$ occurs with multiplicity divisible by r. Since $d^e = 1$ this implies $r = 1$. Hence $G[m]$ is isomorphic to $\rho_f \otimes (\varepsilon\cdot\chi^{k-2})^{-1}$ and $d^0 = 1$.

2) Since $V_pG = T_pG \otimes \mathbb{Q}_p$ is free of rank 2 over $H_m \otimes \mathbb{Q}_p$, the H_m-module T_pG is free (of rank 2) if and only if $T_pG/mT_pG = G[m]$ has dimension 2 over E, by Nakayama's lemma. But $\dim G[m] = d^e + d^0$, so this follows from 1).

3) The ring H_m is local, of dimension 1. It is therefore Gorenstein if and only if the module $H_m^\vee = \mathrm{Hom}(H_m, \mathbb{Z}_p)$ is free of rank 1 [B]. The H_m-module T_pG^e is free of rank 1 over E. Likewise T_pG^0 is free of rank 1, as $d^0 = 1$ by part 1). Hence $T_pG^0 = \mathrm{Hom}(T_pG^e, \mathbb{Z}_p) \simeq \mathrm{Hom}(H_m, \mathbb{Z}_p)$ is free of rank 1, and H_m is Gorenstein.

Note. We do not know if the three equivalent statements in Proposition 12.10 continue to hold when $k = p$ and $a_p^2 = \varepsilon(p)$. The argument using Proposition 11.8 only shows that d^0 is odd.

Using the results in the previous three propositions, we can now complete the proof of Proposition 12.1. If $k = p$ and $a_p^2 = \varepsilon(p)$, these results show that the only character of $\mathscr{G}_p$ which occurs in ρ_f is $\lambda(\varepsilon(p)/a_p) = \lambda(a_p)$, which is the content of 12.1. In all other cases, the filtration $0 \to G^0[m] \to G[m] \to G^e[m] \to 0$ of $\mathscr{G}_p$-modules gives (after twisting by $\varepsilon \cdot \chi^{k-2}$) a filtration $0 \to L \to W \to L' \to 0$ on the representation space W of ρ_f, where $\mathscr{G}_p$ acts on the line L by the character $\lambda(1/a_p) \cdot \chi \cdot \lambda(\varepsilon(p)) \cdot \chi^{k-2} = \lambda(\varepsilon(p)/a_p) \cdot \chi^{k-1}$ and on the line L' by the character $\lambda(a_p/\varepsilon(p)) \cdot \chi^{2-k} \cdot \lambda(\varepsilon(p)) \cdot \chi^{k-2} = \lambda(a_p)$.

§13. Extension classes and companion forms. In this section, we assume that $f = \Sigma a_n q^n$ is a normalized eigenform of type (k, ε) for $\Gamma_1(N)$ with coefficients in the finite field E of characteristic p. We assume the weight k of f satisfies $2 \leqslant k \leqslant p$ and define $k' = p + 1 - k$, so $1 \leqslant k' \leqslant p - 1$. We always assume that $a_p \neq 0$; in the case when $k = p$, we will also assume that $a_p^2 \neq \varepsilon(p)$. Finally, we assume that the representation $\rho_f \colon \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to GL_2(E)$ associated to f is irreducible; when $p \neq 2$ this implies that ρ_f is absolutely irreducible [S8].

By the results of the previous section, we have a realization of the Galois representation $\rho_f \otimes (\varepsilon\chi^{k-2})^{-1}$ on the m-torsion $G[m]$ of the p-divisible group associated to f in the Jacobian. The finite group scheme $G[m]$ over $\mathbb{Q}$ has the structure of an E-vector space scheme of dimension 2; over $\mathbb{Q}_p$ it lies in an exact sequence

$$0 \to G^0[m] \to G[m] \to G^e[m] \to 0 \tag{13.1}$$

of E-vector space schemes, where the E-vector spaces $G^0[m]$ and $G^e[m]$ each have dimension 1. This sequence was the key to our understanding of the restriction of the representation ρ_f to a decomposition group at p in $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$. The E-vector spaces in (13.1) all have canonical flat extensions over $\mathbb{Z}_p[\zeta_p]$. We are now concerned with when the exact sequence in (13.1) is split, and in determining its extension class.

PROPOSITION 13.2. *The following are equivalent:*

1) *The sequence of E-vector space schemes in* (13.1) *is uniquely split over* $\mathbb{Q}_p$.

2) *The sequence of E-vector space schemes over* $\mathbb{Z}_p[\zeta_p]$ *which extends* (13.1) *is uniquely split over* $\mathbb{Z}_p[\zeta_p]$.

3) *The restriction of* ρ_f *to* $\mathscr{G}_p = \mathrm{Gal}(\bar{\mathbb{Q}}_p/\mathbb{Q}_p)$ *is diagonalizable, and is the sum of the distinct characters* $\chi^{k-1}\lambda(\varepsilon(p)/a_p)$ *and* $\lambda(a_p)$.

Proof. The equivalence of 1) and 3) follows from the fact that the representation ρ_f occurs in $G[m] \otimes \varepsilon\chi^{k-2}$. The two characters $\chi^{k-1}\lambda(\varepsilon(p)/a_p)$ and $\lambda(a_p)$ of $\mathscr{G}_p$ on the semi-simplification of ρ_f are given by Proposition 12.1. They are distinct as $\chi^{k-1} = 1$ implies that $k = p$, where we have assumed $a_p^2 \neq \varepsilon(p)$. Hence if (13.1) splits, it splits uniquely.

Part 1) is equivalent to assuming that the sequence (13.1) is uniquely split over the extension $\mathbb{Q}_p(\zeta_p)$, which has degree prime to p over $\mathbb{Q}_p$. This is clearly implied by 2), but it also implies a spliting over $\mathbb{Z}_p[\zeta_p]$. To show this, we may base change to any étale extension R of degree prime to p of $\mathbb{Z}_p[\zeta_p]$, and show that a splitting of (13.1) over the quotient field L of R implies a splitting over R.

Choose R so that the characters $\lambda(1/a_p)$ and $\lambda(a_p/\varepsilon(p))$ are trivial on $\mathrm{Gal}(\bar{\mathbb{Q}}_p/L)$. The E-vector space scheme $G^e[m]$ is isomorphic to the étale vector space scheme $E = E \otimes \mathbb{Z}/p\mathbb{Z}$ with trivial Galois action over R, and $G^0[m]$ is isomorphic to the Cartier dual ${}^tE = E^\vee \otimes \mu_p$ over R, where $\mathrm{E}^\vee = \mathrm{Hom}(E, \mathbb{Z}/p\mathbb{Z})$. The existence of such isomorphisms follows from Propositions 12.8, 12.9. Since Kummer theory gives an isomorphism of $\mathbb{Z}/p\mathbb{Z}$-vector spaces $\mathrm{Ext}^1_R(\mathbb{Z}/p\mathbb{Z}, \mu_p) = H^1(R, \mu_p) = R^*/R^{*^p}$ for any local ring R, where Ext^1_R classifies extensions of finite flat group schemes, we obtain a canonical isomorphism of E vector spaces

$$(13.3) \qquad \mathrm{Ext}_R(E, E^\vee \otimes \mu_p) \xrightarrow{\sim} R^*/R^{*^p} \otimes E^\vee .$$

Here Ext_R classifies extensions in the category of E-vector space schemes: it is the sub-module of the E-bimodule

$$\begin{tikzcd}\mathrm{Ext}^1_R(E, E^\vee \otimes \mu_p) \arrow[d, equal] \\ R^*/R^{*^p} \otimes (E \otimes E^\vee)\end{tikzcd}$$

of extensions of group schemes where the two E actions are the same. Hence the sequence (13.1) over R gives a class in $R^*/R^{*^p} \otimes E^\vee$, which is zero iff the sequence is split. Since R^*/R^{*^p} injects into L^*/L^{*^p}, a splitting over L implies one over R.

We now consider the extension class defined by (13.1) in more detail. By Propositions 12.8 and 12.9 the vector space schemes $G^0[m]$ and $G^e[m]$ are twists of $E^\vee \otimes \mu_p$ and E over $\mathbb{Q}_p$, by the characters $\psi_0 = \lambda(1/a_p)$ and $\psi_e = \chi^{2-k}\lambda(a_p/\varepsilon(p)) = \chi^{k'}\lambda(a_p/\varepsilon(p))$, respectively. In other words, we have isomorphisms $i_0\colon E^\vee \otimes \mu_p \xrightarrow{\sim} G^0[m]$ and $i_e\colon E \xrightarrow{\sim} G^e[m]$ over L such that $i_0^\sigma = i_0 \cdot \psi_0(\sigma)$, $i_e^\sigma = i_e \cdot \psi_e(\sigma)$, for all $\sigma \in \mathrm{Gal}(L/\mathbb{Q}_p)$. Let $\alpha\colon G^0[m] \to E^\vee \otimes \mu_p$ and $\beta\colon E \to G^e[m]$ be homomorphisms of flat E-vector space schemes over R. Via push-out and pull-back we obtain a homomorphism $\alpha_*\beta^*\colon \mathrm{Ext}_R(G^e[m], G^0[m]) \to \mathrm{Ext}_R(E, E^\vee \otimes \mu_p)$. By (13.3), the extension $G[m]$ defines a class

$$(13.4) \qquad q_m(\alpha, \beta) \quad \text{in } R^*/R^{*^p} \otimes E^\vee .$$

We will write the value group $R^*/R^{*^p} \otimes E^\vee$ multiplicatively, for reasons that will become clear in the next section.

Proposition 13.5. 1) *For s and t in E, we have the formula $q_m(s\alpha, t\beta) = q_m(\alpha, \beta)^{st}$. Hence the classes $q_m(\alpha, \beta)$ all lie in an E-subspace $\langle q_m \rangle$ of dimension ≤ 1 in $R^*/R^{*^p} \otimes E^\vee$.*

2) *For* $\sigma \in \mathrm{Gal}(L/\mathbb{Q}_p)$, *we have the formula* $q_m(\alpha, \beta)^\sigma = q_m(\alpha, \beta)^{\psi(\sigma)}$, *where* ψ *is the character* $\chi^{k'} \cdot \lambda(a_p^2/\varepsilon(p))$. *Hence* $\langle q_m \rangle$ *is contained in the* ψ*-eigenspace of* $R^*/R^{*p} \otimes E^\vee$.

3) *We have* $\langle q_m \rangle = 1$ (*i.e.*, $q_m(\alpha, \beta) \equiv 1 \pmod{R^{*p}}$) *if and only if the exact sequence of E-vector space schemes in* (13.1) *is split.*

Proof. 1) is clear, as is the formula $q_m(\alpha, \beta)^\sigma = q_m(\alpha^\sigma, \beta^\sigma)$. Taking $\alpha = i_0^{-1}$ and $\beta = i_e$ shows that $\langle q_m \rangle$ lies in the $\psi = \psi_e/\psi_0$-eigenspace, which proves 2), and 3) is a tautology from the definition of Ext groups.

The ring R is an unramified extension of $\mathbb{Z}_p[\zeta_p]$, so $\pi = (1 - \zeta_p)$ is a uniformizing parameter in R. Since $R^*/(1 + \pi R)$ has order prime to p, $R^*/R^{*p} = (1 + \pi R)/(1 + \pi R)^p$. Since $(1 + \pi a)^p = 1 + p\pi a + \cdots + \pi^p a^p$, and π^{p-1} has the same valuation as p, we see that $(1 + \pi R)^p$ is contained in the subgroup $(1 + p\pi R) = 1 + \pi^p R$. Hence for $i = 1, 2, 3, \ldots, p$ we have the quotient $1 + \pi^i R/(1 + \pi R)^p = U_i$, and a descending filtration

$$R^*/R^{*p} = U_1 \supset U_2 \supset U_3 \supset \cdots \supset U_p \supset 0. \tag{13.6}$$

The subgroups U_i are all stable under the action of $\mathrm{Gal}(L/\mathbb{Q}_p)$, where L is the quotient field of R. The group U_p is isomorphic to $\mathbb{Z}/p\mathbb{Z}$ and $\mathrm{Gal}(L/\mathbb{Q}_p)$ acts via the character χ. For $i = 1, 2, \ldots, p - 1$ the representation of $\mathrm{Gal}(L/\mathbb{Q}_p)$ on U_i/U_{i+1} is induced from the character χ^i of the inertia subgroup. In particular, the $\chi^{k'} \cdot \lambda(a_p^2/\varepsilon(p))$-eigenspace of $R^*/R^{*p} \otimes E^\vee$ has dimension 1 over E, since $a_p^2 \neq \varepsilon(p)$ when $k = p$.

PROPOSITION 13.7. *The three conditions of Proposition* 13.2 *are equivalent to the following*:

1) *When f has weight* $2 \leqslant k \leqslant p - 1$, *the representation* ρ_f *is tamely ramified at p.*

2) *When f has weight* $k = p$, *the representation* ρ_f *is unramified at p.*

Proof. Clearly part 3) of Proposition 13.2 implies that ρ_f is tamely ramified at p when $2 \leqslant k \leqslant p - 1$, and unramified at p when $k = p$. Conversely, the restriction of ρ_f to the inertia subgroup of $\mathrm{Gal}(\bar{\mathbb{Q}}_p/\mathbb{Q}_p(\zeta_p))$ has matrix form:

$$\begin{pmatrix} 1 & \eta \\ 0 & 1 \end{pmatrix}$$

where η is a character with values in E. If ρ_f is tamely ramified or unramified at p, the character η is trivial and the sequence (13.1) is split over the maximal unramified extension of $\mathbb{Q}_p(\zeta_p)$. Hence the extension class $\langle q_m \rangle$ lies in the subspace $U_p \otimes E^\vee$ of $R^*/R^{*p} \otimes E^\vee$. But it lies in the eigenspace with character $\chi^{k'} \cdot \lambda(a_p^2/\varepsilon(p)) \neq \chi$, so $\langle q_m \rangle \equiv 1$ and the extension (13.1) is split over $\mathbb{Q}_p$.

The following modular criterion for the restriction of ρ_f to $\mathscr{G}_p$ to be diagonalizable was noted by Serre [S7, pg. 18].

PROPOSITION 13.8. *If $f = \Sigma a_n q^n$ has weight k with $2 \leqslant k \leqslant p - 1$, assume that there is a cusp form $g = \Sigma b_n q^n$ of weight $k' = p + 1 - k$ and character ε for $\Gamma_1(N)$ over E which satisfies: $\theta g = \theta^{k'} f$ in $\tilde{M}_{k'+p+1}$ and $g | U_p = \lambda \cdot g$. If f has weight $k = p$ assume that there is a cusp form g of weight p and character ε for $\Gamma_1(N)$ over E which satisfies: $\theta g = \theta f$ in $\tilde{M}_{p+2}$, $g \neq f$, and $g | U_p = \lambda g$. Then*

1) *The cusp form g is a normalized eigenform for $\Gamma_1(N)$, and the eigenvalues b_l for T_l and U_l satisfy $b_l = a_l \cdot l^{k'-1}$.*

2) *The eigenvalue $\lambda = b_p$ of U_p is nonzero, and is given by the formula $b_p = \varepsilon(p)/a_p$.*

3) *The representation ρ_f is diagonalizable when restricted to $\mathscr{G}_p$: all the equivalent conditions in Propositions 13.2 and 13.7 hold.*

Proof. 1) The identity $\theta g = \theta^{k'} f$ shows that $b_1 = 1$. To show that $g | T_l = a_l \cdot l^{k'-1} g$, we use formula (4.8):

$$l \cdot \theta(g | T_l) = \theta g | T_l = (\theta^{k'} f) | T_l = l^{k'} \theta^{k'} (f | T_l).$$

But $f | T_l = a_l f$, so $\theta(g | T_l) = \theta(a_l \cdot l^{k'-1} g)$. Since the kernel of θ is the image of V_p, we have

$$g | T_l - a_l \cdot l^{k'-1} g = h | V_p. \tag{13.9}$$

When $k \neq p$ this forces $h = 0$, as the left hand side of (13.9) has weight $k' < p$. When $k = p$, h is a form of weight 1. But the left hand side of (13.9) is, by hypothesis, an eigenvector for U_p with eigenvalue λ. Since $(h | V_p) | U_p = h$, this shows that $h = \lambda \cdot (h | V_p)$ which forces $h = 0$. A similar proof works for U_l.

2) If $k \neq p$ and $\lambda = 0$, we have $g = \theta^{k'-1} g$ in $\tilde{M}_{k'}$ by part 1). This contradicts the fact that $\theta^{k'-1} f$ has filtration pk', which was noted in the proof of Proposition 4.10. If $k = p$ and $\lambda = 0$, we have $g = \theta^{p-1} f$ in $\tilde{M}_p$, by part 1). This contradicts the fact that $\theta^{p-1} f$ has filtration p^2, which was proved in Proposition 4.10, part 2).

Let ρ_g be the Galois representation associated to the normalized eigenform g; by Proposition 12.1 there is a line in the semi-simplification of $\rho_g | \mathscr{G}_p$ where the local Galois group acts by the character $\lambda(b_p)$. But part 1) implies that $\rho_g \simeq \rho_f \otimes \chi^{k'-1}$ as representations of $\mathscr{G} = \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$, and the characters of $\mathscr{G}_p$ in the semi-simplification of $\rho_f \otimes \chi^{k'-1}$ are $\chi^{k'-1} \cdot \lambda(a_p)$ and $\lambda(\varepsilon(p)/a_p)$, respectively. When $k \neq p$ the first character $\chi^{k'-1} \lambda(a_p)$ is ramified, so we must have $b_p = \varepsilon(p)/a_p$. When $k = p$, both character are unramified, but the hypothesis $g \neq f$ means that $b_p \neq a_p$. Hence $b_p = \varepsilon(p)/a_p$ as claimed.

3) By Proposition 12.1 the representation space W of $\rho_f \simeq \rho_g \otimes \chi^{k-1}$ has the $\mathscr{G}_p$-stable lines L_f and L_g with characters $\chi^{k-1} \lambda(\varepsilon(p)/a_p)$ and $\chi^{k-1}(\chi^{k'-1} \cdot \lambda(\varepsilon(p)/b_p)) = \lambda(a_p)$, respectively. Since $a_p^2 \neq \varepsilon(p)$ when $k = p$, these two lines are distinct and $W = L_f \oplus L_g$ is semi-simple as a $\mathscr{G}_p$-module.

Serre calls the modular form g (when it exists) a companion form of f. This relation is clearly symmetric: f is a companion form of g. Proposition 13.8 asserts that forms

f which have a companion g have locally split Galois representations at p. In the next three sections, we will prove the following converse to Proposition 13.8, which was conjectured by Serre [S7, pg. 18].

THEOREM 13.10. *Let $f = \Sigma a_n q^n$ be a normalized eigenform of type (k, ε) for $\Gamma_1(N)$ over E, with $2 \leqslant k \leqslant p$* and *$a_p \neq 0$. When $k = p$ we assume further that $a_p^2 \neq \varepsilon(p)$.*

If the restriction of the Galois representation $\rho_f: \mathcal{G} \to GL_2(E)$ to $\mathcal{G}_p$ is diagonalizable, then a companion form $g = \Sigma b_n q^n$ exists, satisfying the hypotheses of Proposition 13.8.

COROLLARY 13.11. *Let $f = \Sigma a_n q^n$ be a normalized eigenform of type (k, ε) for $\Gamma_1(N)$ over E, with $2 \leqslant k \leqslant p$ and $a_p \neq 0$. When $k = p$ we assume further that $a_p^2 \neq \varepsilon(p)$. Then the following are equivalent:*

1) *The Galois representation ρ_f is tamely ramified $(k \neq p)$ or unramified $(k = p)$ at p.*

2) *There is a form $h = \Sigma c_n q^n$ of type (k', ε) for $\Gamma_1(N)$ over E, with $k' = p + 1 - k$, which satisfies the differential equation $\theta h = \theta^{k'} f$ in $\tilde{M}_{k'+p+1}$.*

Proof of Corollary. If ρ_f is a reducible representation of $\mathcal{G}$, it is completely reducible (as ρ_f is semi-simple). In this case, ρ_f is certainly diagonalizable when restricted to $\mathcal{G}_p$ and 1) always holds. But one can show, using the theory of Eisenstein series, that a form h satisfying 2) also exists. For example, if

$$f = \frac{1}{2}L(1-k, \varepsilon) + \sum_{n \geqslant 1}\left(\sum_{d|n} \varepsilon(d) d^{k-1}\right) q^n$$

is the Eisenstein series of weight $2 \leqslant k \leqslant p$ and character $\varepsilon \neq 1$, then $\rho_f = 1 \oplus \varepsilon\chi^{k-1}$ and the form h is given by the Eisenstein series

$$h = \sum_{n \geqslant 1}\left(\sum_{d|n} \varepsilon(n/d) d^{k'-1}\right) q^n$$

of weight k' and character ε. The general case is similar, but we do not treat it here. Henceforth, we assume ρ_f is irreducible as a representation of $\mathcal{G}$.

First assume that $k \neq p$. By Proposition 13.7 and Theorem 13.10 we have 1) $\Rightarrow$ 2): simply take $h = g$, the companion form. Conversely, if h exists and satisfies $\theta h = \theta^{k'} f$, the proof of 13.8 shows that h is an eigenvector for T_ℓ and U_ℓ. We may find an eigenvector g for U_p in the span of $\langle h, h|U_p, h|U_p^2 \ldots \rangle$; since this has the same eigenvalues for T_ℓ and U_ℓ as h, it is a normalized eigenform which satisfies $\theta g = \theta^{k'} f$. Hence g is a companion to f, and the restriction of ρ_f to $\mathcal{G}_p$ is diagonalizable by Proposition 13.8.

Now assume $k = p$. If ρ_f is unramified at p, then by Proposition 13.7 and Theorem 13.10 we have a companion g to f of weight p which satisfies $g|U_p = b_p g$ with $b_p = \varepsilon(p)/a_p \neq a_p$. We define the form h of weight $k' = 1$ by the identity $(a_p - b_p) \cdot h|V_p = f - g$. Conversely, if h exists, we may assume it is a normalized eigenform of weight 1 (it is an eigenvector for T_ℓ and U_l from $\theta h = \theta f$, and one can choose an eigenvector for T_p in the span of $\langle h, h|T_p, h|T_p^2, \ldots\rangle$). Let g be the eigenvector for U_p with eigenvalue $b_p = \varepsilon(p)/a_p$ in the span of $\langle h|V_p, Ah\rangle$. Then g is a companion for f, and ρ_f is unramified at p by Proposition 13.8.

We note that the identity $\theta h = \theta^{k'}f = f'$ of part 2) of Corollary 13.11 can be verified by a finite amount of computation. It suffices to check that $nc_n = n^{k'}a_n$ for all $n \leqslant \frac{1}{24}N^2\Pi_{q|N}(1-q^{-2})\cdot(k'+p+1)$, as any form of weight $(k'+p+1)$ on $\Gamma_1(N)$ which vanishes to this order at ∞ must be identically zero.

R. Coleman has observed that the existence of a companion form for f can be neatly expressed in terms of the vanishing of the class of the meromorphic differential $v_{f'}$ in the de Rham cohomology of the Igusa curve $I = I_1(N)$ [Col]. Recall that $f' = \theta^{k'}f$ has filtration $k'+p+1$, so $v_{f'} = f'(q)dq/q$ is a meromorphic exact differential on I which is regular outside Σ, and has poles of order $\leqslant k'+1$ of supersingular points.

The de Rham cohomology $H^1(I/E)$ of the Igusa curve I over E is defined as the first hypercohomology group of the complex $\Omega_{I/E} = (\mathcal{O}_{I/E} \xrightarrow{d} \Omega^1_{I/E})$. Coleman [Co2] shows that $H^1(I/E)$ is isomorphic to the quotient of the space of meromorphic differentials v on X, with no residues and poles of order $\leqslant p$ at all points x, by the space of exact differentials dg, where the function g has poles of order $\leqslant(p-1)$ at all points x. Hence the differentials v_f and $v_{f'}$ define classes in $H^1(I/E)$, where E is the finite field generated by the coefficients of f, and their cup product $\langle v_f, v_{f'}\rangle$ lies in E. Since v_f is holomorphic, and the poles of $v_{f'}$ are contained in Σ, the cup product is given by the formula

$$\langle v_f, v_{f'}\rangle = -\sum_{x\in\Sigma} Res_x(g_x\cdot v_f) \tag{13.12}$$

where, for each $x \in \Sigma$, g_x is a meromorphic function on I such that the differential $v_{f'} - dg_x$ is regular at x.

The expansions of v_f and $v_{f'}$ at supersingular points x was determined in Proposition 9.9. If z is a local parameter at x we have

$$\begin{cases} v_f = \left(\alpha_x\cdot z^{p-k} + \displaystyle\sum_{n\geqslant p-1}\alpha_n z^n\right)dz \\[2ex] v_{f'} = \left(\beta_x\cdot z^{-1-k'} + \displaystyle\sum_{n\geqslant 0}\beta_n z^n\right)dz. \end{cases}$$

Hence $g_x = (-1/k')\cdot\beta_x\cdot z^{-k'} + \Sigma_{n\geqslant 0}\gamma_n z^n$ and

$$\begin{cases} Res_x(g_x\cdot v_f) = \dfrac{-\alpha_x\cdot\beta_x}{k'}, \\[2ex] k'\langle v_f, v_{f'}\rangle = \displaystyle\sum_{x\in\Sigma}\alpha_x\beta_x \quad \text{in } E. \end{cases} \tag{13.13}$$

PROPOSITION 13.14. *The following are equivalent*:

1) *The class $v_{f'}$ is zero in $H^1(I/E)$*
2) *There is a modular form h of weight k' with $\theta h = f'$*
3) *The representation ρ_f is diagonalizable when restricted to $\mathscr{G}_p$.*

Assuming H_m is unramified over $\mathbb{Z}_p$, these three conditions are equivalent to:

4) *The cup-product $\langle v_f, v_{f'} \rangle = 0$ in E.*

Proof. We have $v_{f'} \equiv 0$ in $H^1(I/E)$ if and only if there is a function h on I which satisfies $dh = v_{f'}$ *and* h has poles of order $\leqslant (p-1)$ at all supersingular x. Then h corresponds to a form of weight k' satisfying $\theta h = f'$, so 1) $\Leftrightarrow$ 2). We have seen that 2) $\Leftrightarrow$ 3) in Corollary 13.11. Clearly 1) $\Rightarrow$ 4) in all cases. When H_m is unramified over $\mathbb{Z}_p$, there is a unique lifting F of f, in the sense of §9, to an eigenform of weight 2 in characteristic zero. Hence the eigenvalues of f do not occur in the Hecke module $M_k/\langle f \rangle$ by the multiplicity one theorem. Since the cup-product pairing induces a non-degenerate duality between the f- and f'-eigenspaces for the Hecke algebra on $H^1(I/E)$, we have $v_{f'} = 0$ if and only if $\langle v_f, v_{f'} \rangle = 0$.

Proposition 13.14 suggests that there may be a formula for $\langle v_f, v_{f'} \rangle$ in terms of the invariant q_m of the extension of E-vector spaces

$$0 \to G^0[m] \to G[m] \to G^e[m] \to 0.$$

§14. The invariant of Serre and Tate. In this section, we refine the invariant $q_m(\alpha, \beta)$ which classifies the extension $0 \to G^0[m] \to G[m] \to G^e[m] \to 0$ of E-vector space schemes over $\mathbb{Q}_p$ by studying the extension class of the sequence $0 \to G^0 \to G \to G^e \to 0$ of p-divisible groups with endomorphisms by H_m.

Let R be a complete, discrete valuation ring with quotient field K of characteristic zero and perfect residue field of characteristic p. We say a p-divisible group A over R is m-divisible of height g if A has endomorphisms by H_m and T_pA is a free H_m-module of rank g; the height of A as a p-divisible group is then equal to $g \cdot h = g \cdot [H_m : \mathbb{Z}_p]$. By Tate's theorem [T], the functor $A \quad T_pA$ embeds the category of m-divisible groups over R as a full sub-category of the category of free H_m-modules of finite rank with a continuous, H_m-linear action of $\mathrm{Gal}(\bar{K}/K)$. The groups $\mathrm{Ext}_R(A, B)$ relative to the category of m-divisible groups classify extensions $0 \to B \to C \to A \to 0$ with C m-divisible over R; $\mathrm{Ext}_R(A, B)$ is the subgroup of the bi-H_m-module $\mathrm{Ext}_R^1(A, B)$ which classifies extensions of p-divisible groups where the two H_m-actions are the same. In particular, $\mathrm{Ext}_R(A, B)$ has the natural structure of an H_m-module.

We apply this to the p-divisible group G associated to the eigenform $f = \Sigma a_n q^n$ of weight $2 \leqslant k \leqslant p$, with $a_p \neq 0$ and $a_p^2 \neq \varepsilon(p)$ when $k = p$. By Proposition 12.10 the groups G, G^0, and G^e are m-divisible over $\mathbb{Q}_p$ or $\mathbb{Z}_p[\zeta_p]$. Let R be the completion of the maximal unramified extension of $\mathbb{Z}_p[\zeta_p]$, and let $\mathbb{Q}_p/\mathbb{Z}_p \otimes H_m$ be the trivial étale m-divisible group of height 1 over R, i.e., $T_p(\mathbb{Q}_p/\mathbb{Z}_p \otimes H_m) = H_m$ with trivial Galois action. The Cartier dual of the trivial étale m-divisible group, in the category of p-divisible groups over R, is the multiplicative p-divisible group $\hat{\mathbb{G}}_m \otimes H_m^\vee =$

$\mathrm{Hom}_{\mathbb{Z}_p}(H_m, \hat{\mathbb{G}}_m)$. Since $H_m^\vee = \mathrm{Hom}(H_m, \mathbb{Z}_p)$ is a free H_m-module of rank 1 by Proposition 12.10 (H_m is Gorenstein), $T_p(\hat{\mathbb{G}}_m \otimes H_m^\vee) = T_p(\hat{\mathbb{G}}_m) \otimes H_m^\vee$ is free of rank 1 over H_m and the group $\hat{\mathbb{G}}_m \otimes H_m^\vee$ is also m-divisible of height 1 over R. By part 3) of Proposition 12.9, we have isomorphism of m-divisible groups over R:

$$\begin{cases} j_e\colon \mathbb{Q}_p/\mathbb{Z}_p \otimes H_m \stackrel{\sim}{\rightarrow} G^e \\ j_0\colon \hat{\mathbb{G}}_m \otimes H_m^\vee \stackrel{\sim}{\rightarrow} G^0 \end{cases}$$

such that $j_e^\sigma = j_e \cdot \eta_e(\sigma)$, $j_0^\sigma = j_0 \cdot \eta_0(\sigma)$ for all $\sigma \in \mathrm{Gal}(K/\mathbb{Q}_p)$. Here K is the quotient field of R and the characters η_e and η_0 take values in H_m^* and reduce to the characters ψ_e and ψ_0 (modulo m), which describe the twisting of $G^e[m]$ and $G^0[m]$ respectively. When $3 \leqslant k \leqslant p$ we have $\eta_0 = \lambda(U_p^{-1})$ and $\eta_e = \lambda(U_p \cdot \langle p \rangle_N^{-1})\chi^{k'}$. When $k = 2$ we have $\eta_0 = \lambda(u^{-1})$ and $\eta_e = \lambda(u \cdot \langle p \rangle_N^{-1})$ where u is the unit root of $x^2 - T_p x + \langle p \rangle_N \cdot p = 0$ in H_m.

Since R is complete and local, we have a canonical isomorphism:

$$\begin{aligned} \text{(14.1)} \qquad \mathrm{Ext}_R^1(\mathbb{Q}_p/\mathbb{Z}_p, \hat{\mathbb{G}}_m) &= \varprojlim_n \mathrm{Ext}_R^1(\mathbb{Z}/p^n\mathbb{Z}, \mu_{p^n}) \\ &= \varprojlim_n H^1(R, \mu_{p^n}) \\ &= \varprojlim_n R^*/R^{*^{p^n}} \\ &= 1 + \pi R \end{aligned}$$

where $\pi = (1 - \zeta_p)$ is a uniformizing parameter in R. The group $(1 + \pi R)$ is a $\mathbb{Z}_p$-module, and we write this action multiplicatively. From (14.1) it follows that:

$$\text{(14.2)} \qquad \mathrm{Ext}_R(\mathbb{Q}_p/\mathbb{Z}_p \otimes H_m, \hat{\mathbb{G}}_m \otimes H_m^\vee) = (1 + \pi R) \otimes_{\mathbb{Z}_p} H_m^\vee .$$

Now let $\alpha\colon \mathbb{Q}_p/\mathbb{Z}_p \otimes H_m \to G^e$ and $\beta\colon G^0 \to \hat{\mathbb{G}}_m \otimes H_m^\vee$ be homomorphisms of m-divisible groups over R. By push-out and pull-back they give a homomorphism $\alpha^*\beta_*$ from $\mathrm{Ext}_R(G^e, G^0)$ to $\mathrm{Ext}_R(\mathbb{Q}_p/\mathbb{Z}_p \otimes H_m, \hat{\mathbb{G}}_m \otimes H_m^\vee)$. Applying the homomorphism to the class of the extension $1 \to G^0 \to G \to G^e \to 1$ over R, and using (14.2) gives an extension class:

$$\text{(14.3)} \qquad q(\alpha, \beta) \quad \text{in} \quad (1 + \pi R) \otimes H_m^\vee .$$

PROPOSITION 14.4. 1) *For all s, t in H_m we have $q(s\alpha, t\beta) = q(\alpha, \beta)^{st}$.*

2) *For all $\sigma \in \mathrm{Gal}(K/\mathbb{Q}_p)$ we have $q(\alpha, \beta)^\sigma = q(\alpha, \beta)^{\eta(\sigma)}$, where $\eta = \eta_e/\eta_0 = \lambda(u^2 \cdot \langle p \rangle_N^{-1}) \cdot \chi^{k'}$ and u is a unit in H_m with $u \equiv a_p$ (modulo m).*

3) *The invariants $q(\alpha, \beta)$ of the extension defined by G lie on the H_m-submodule of $(1 + \pi R) \otimes H_m^\vee$ where $\mathrm{Gal}(K/\mathbb{Q}_p)$ acts via the character η. In particular, $q(\alpha, \beta)$ is contained in the H_m-submodule $(1 + \pi^{k'} R) \otimes H_m^\vee$.*

4) *The image of* $q(\alpha, \beta)$ *in the quotient* $(1 + \pi R) \otimes H_m^\vee/(1 + \pi R) \otimes mH_m^\vee$ *lies on the E-subspace* $\langle q_m \rangle$ *of* $R^*/R^{*^p} \otimes E^\vee \approx R^*/R^{*^p} \otimes E$.

Proof. This is similar to Proposition 13.5, and we leave the details to the reader. When H_m is unramified over $\mathbb{Z}_p$, the quotient $H_m^\vee/mH_m^\vee$ is canonically isomorphic to $E^\vee$, and the image of $q(\alpha, \beta)$ in the quotient is equal to $q_m(\alpha, \beta)$. When H_m is ramified, the isomorphism is not canonical and $H_m^\vee/mH_m^\vee$ and $E^\vee$ are simply two 1-dimensional E-vector spaces. Similarly $G^e[m]$ is isomorphic to $m^{-1}pH_m/pH_m$ rather than to H_m/mH_m.

The choice of a homomorphism $\alpha\colon \mathbb{Q}_p/\mathbb{Z}_p \otimes H_m \to G^e$ of m-divisible groups is equivalent to the choice of an element $\bar{\alpha} \in T_p(G^e) = T_p(\bar{G})$. Similarly, the choice of a homomorphism $\beta\colon G^0 \to \hat{\mathbb{G}}_m \otimes H_m^\vee$ is equivalent to the choice of an element $\bar{\beta} \in \mathrm{Hom}(T_pG^0, \mathbb{Z}_p(1)) \xrightarrow{\sim} T_p(({}^tG)^e) = T_p({}^t\bar{G})$. If we set $q_G(\bar{\alpha}, \bar{\beta}) = q(\alpha, \beta)$ we obtain, by Proposition 14.4 an H_m-bilinear pairing

$$(14.5) \qquad q_G\colon T_p\bar{G} \times T_p{}^t\bar{G} \to (1 + \pi R) \otimes H_m^\vee$$

with image in the η-eigenspace for $\mathrm{Gal}(K/\mathbb{Q}_p)$. This is the Serre-Tate invariant of the m-divisible group G: it takes values in $(1 + \pi^{k'}R) \otimes mH_m^\vee$ if and only if the exact sequence of E-vector space schemes $0 \to G^0[m] \to G[m] \to G^e[m] \to 0$ is split over R.

The usual Serre-Tate invariant (cf. [K5]) of the ordinary p-divisible group G over R is the $\mathbb{Z}_p$-bilinear pairing

$$q_G'\colon T_p\bar{G} \times T_p{}^t\bar{G} \to (1 + \pi R).$$

This satisfies $q_G'(h\alpha, \beta) = q_G'(\alpha, h\beta)$ for all $h \in H_m$; the action of th on ${}^tG$ is identified with the action of h on G in the isomorphism of G with its dual (cf. the definition of $\langle\ ,\ \rangle$ in the proof of Proposition 11.4). If $\mathrm{tr}\colon H_m^\vee \to \mathbb{Z}_p$ is the canonical $\mathbb{Z}_p$-linear map $\mathrm{tr}(f) = f(1)$, we have

$$(14.6) \qquad q_G' = (1 \otimes \mathrm{tr})q_G.$$

Indeed, tr arises from the contraction $H_m \otimes H_m^\vee \to \mathbb{Z}_p$ which forgets the m-divisible structure on G.

Since $ord_\pi(p) = p - 1$, the ideal πR has divided powers (if $x \in \pi R$ then $x^n/n! \in \pi R$) and the p-adic logarithm defined by the series $\log(1 + x) = x - x^2/2 + x^3/3 - x^4 \ldots$ gives a $\mathbb{Z}_p$-linear continuous homomorphism $\log\colon (1 + \pi R) \to \pi R$. Since the ideal π^2R has topologically nilpotent divided powers ($x^n/n! \to 0$ as $n \to \infty$), the p-adic logarithm induces an isomorphism $1 + \pi^iR \xrightarrow{\sim} \pi^iR$ for all $i \geqslant 2$, its inverse is defined by the series $\exp(y) = 1 + y + y^2/2! + y^3/3! + \cdots$. It follows that the homomorphism $\log\colon 1 + \pi R \to \pi R$ is surjective, with kernel the torsion subgroup μ_p of $1 + \pi R$; it induces the Artin Schreier isogeny $t \mapsto t^p - t$ from $1 + \pi R/1 + \pi^2R$ to $\pi R/\pi^2R$.

We define $\log q_G$ as the composite of q_G with the H_m-linear map $\log\otimes 1\colon (1+\pi R)\otimes H_m^\vee \to \pi R\otimes H_m^\vee$.

PROPOSITION 14.7. *The bilinear pairing of $R\otimes H_m$-modules*:

$$\log q_G\colon (R\otimes T_p\bar{G})\times(R\otimes T_p{}^t\bar{G})\to \pi R\otimes H_m^\vee$$

takes values in $\pi^{k'}R\otimes H_m^\vee$, and takes values in the submodule $\pi^{k'}R\otimes mH_m^\vee$ if and only if the sequence of E-vector spaces $0\to G^0[m]\to G[m]\to G^e[m]\to 0$ is split over R.

Proof. We must verify that the η-eigenspace of $\mathrm{Gal}(K/\mathbb{Q}_p)$ containing the image of q_G has no intersection with the kernel of the p-adic logarithm. But the kernel μ_p lies in the χ-eigenspace, and $\eta\not\equiv\chi$ (modulo m) (as $a_p^2\neq\varepsilon(p)$ when $k=p$).

Similarly, we define the pairing of R-modules

$$\log q_G'\colon (R\otimes T_p\bar{G})\times(R\otimes T_p{}^t\bar{G})\to \pi R \tag{14.8}$$

as the composite of q_G' with log: $1+\pi R\to\pi R$. Then $\log q_G'=(\log\otimes\mathrm{tr})q_G$. In the next section, we will present Dwork's formula for $\log q_G'$, using the action of Frobenius on the deRham cohomology of G.

§15. de Rham cohomology and Dwork's formula. We begin with the case when f has weight $k=2$. In this case G is a p-divisible subgroup of the Jacobian $J=J_1(N)$ of the curve $X=X_1(N)$ over $\mathbb{Z}_p$. Since X and J are smooth and proper over $\mathbb{Z}_p$, the de Rham cohomology groups $H^1(X)=\mathbb{H}^1(X,\Omega_X^{\cdot})$ and $H^1(J)=\mathbb{H}^1(J,\Omega_J^{\cdot})$ are free $\mathbb{Z}_p$-modules, and pull-back via the Albanese map $X\hookrightarrow J$ induces an isomorphism $H^1(J)\overset{\sim}{\to}H^1(X)$ [K4]. In the exact sequences

$$\begin{cases} 0\to H^0(X,\Omega_X^1)\to H^1(X)\to H^1(X,\mathcal{O}_X)\to 0\\ 0\to \qquad \omega_J \qquad \to H^1(J)\to \quad \mathrm{Lie}({}^tJ)\quad \to 0\end{cases}\tag{15.1}$$

the invariant differentials ω_J are identified with the holomorphic differentials on X and the Lie algebra of ${}^tJ=\mathrm{Pic}^0(X)$ is identified with $H^1(X,\mathcal{O}_X)$.

The Hecke algebra H_p of $X_1(N)$ acts $\mathbb{Z}_p$-linearly on $H^1(X)$ and preserves the sub-module $H^0(X,\Omega_X^1)$. We define

$$\left\{\begin{matrix} H^1(G)=\varepsilon_m\cdot H^1(X)=\varepsilon_m H^1(J)\\ \omega_G=\varepsilon_m\cdot H^0(X,\Omega^1)=\varepsilon_m\omega_J\end{matrix}\right\}\tag{15.2}$$

These are both H_m-modules, and we have an exact sequence:

$$0\to\omega_G\to H^1(G)\to\mathrm{Lie}({}^tG)\to 0.\tag{15.3}$$

Let $\bar{J}$ be the reduction of J over $\mathbb{Z}/p\mathbb{Z}$, and let $D(\bar{J})$ be the Dieudonné module of the p-divisible group of $\bar{J}$, with its $\mathbb{Z}_p$-linear action of $\mathrm{F}=\mathrm{Fr}_p$ and $\mathrm{V}=\mathrm{Ver}_p$. The ideal $p\mathbb{Z}_p$ has divided powers, so by a theorem of Grothendieck (cf. [G1], [G2]) there is a canonical isomorphism $H^1(J) \xrightarrow{\sim} D(\bar{J})$ which presumably commutes with the action of H_p. (This may be known, but I could not find a precise reference.) Thus the $\mathbb{Z}_p$-module $H^1(J)$ has a filtration, defined by (15.1), as well as a $\mathbb{Z}_p$-linear action of Fr_p. It is extremely difficult to construct the action of Fr_p on $H^1(J)$ directly, or equivalently, to describe the submodule ω_J of $D(\bar{J})$. (See however [F1] for the case of the Dieudonné module of a formal group.) Some general references on Grothendieck's isomorphism, and its relation to crystalline cohomology, are [BM], [BBM], [Br], [K4] and [Me]. Passing to ε_m-eigencomponents, we obtain an isomorphism of H_m-modules

$$(15.4) \qquad H^1(G) \xrightarrow{\sim} D(\bar{G}) = U \oplus P.$$

Here $U = D(\bar{G}^e)$ is the unit root eigenspace for Fr_p and $P = D(\bar{G}^m)$ is the p-root eigenspace for Fr_p. Since Fr_p annihilates the submodule $\omega_G/p\omega_G$ in $H^1(G)/pH^1(G)$, the H_m-submodule ω_G is disjoint from U. Hence, projection onto the second factor gives an isomorphism $\omega_G \xrightarrow{\sim} P$. Since $U = \mathrm{Hom}(T_p\bar{G}, \mathbb{Z}_p)$ and $P = T_p{}^t\bar{G}$, Proposition 12.10 shows that U, P and ω_G are free H_m-modules of rank 1, and consequently that $H^1(G)$ is a free H_m-module of rank 2.

We now consider the case when f has weight $3 \leqslant k \leqslant p$, where G appears as a p-divisible subgroup of $J_1(Np)$. Here we let X be the regular model for $X_1(Np)$ over $\mathbb{Z}_p[\zeta_p]$ studied in §7 and let J be the Néron model of $J_1(Np)$ over this base. We have an isomorphism $\mathrm{Pic}^0(X) \xrightarrow{\sim} {}^tJ^0$ of smooth group schemes over $\mathbb{Z}_p[\zeta_p]$, where ${}^tJ^0$ is the connected component of the Néron model of the dual abelian variety ${}^tJ$ [Ra]. Let $\Omega_X = \Omega_{X/\mathbb{Z}_p[\zeta_p]}$ be the dualizing sheaf on X and define $H^1(X) = \mathbb{H}^1(X, \mathcal{O}_X \xrightarrow{d} \Omega_X)$. Since the differentials $d\colon H^i(X, \mathcal{O}_X) \to H^i(X, \Omega_X)$ are all zero, the spectral sequence for hypercohomology degenerates at the E_1 term and we have an exact sequence of free $\mathbb{Z}_p[\zeta_p]$-modules

$$0 \to H^0(X, \Omega_X) \to H^1(X) \to H^1(X, \mathcal{O}_X) \to 0.$$

The lattice $H^1(X)$ is self-dual with respect to the cup product on $H^1(X) \otimes \mathbb{Q}_p(\zeta_p)$, which is the de Rham cohomology of $X_1(Np)$. We define $H^1(J)$ as the Lie algebra of the smooth group scheme $\mathrm{Extrig}(J^0, \mathbb{G}_m)$ (which represents extensions $0 \to \mathbb{G}_m \to E \to J^0 \to 0$ together with an invariant differential on E pulling back to dt/t, cf. [MM]). As $\mathrm{Ext}(J^0, \mathbb{G}_m) = {}^tJ$ we have an exact sequence of Lie algebras:

$$0 \to \omega_J \to H^1(J) \to \mathrm{Lie}({}^tJ) \to 0$$

as in (15.1). The isomorphism of de Rham cohomology

$$H^1(J) \otimes \mathbb{Q}_p(\zeta_p) \xrightarrow{\sim} H^1(X) \otimes \mathbb{Q}_p(\zeta_p)$$

induced by the Albanese map identifies the lattices $H^1(J)$ and $H^1(X)$, as well as the submodules ω_J and $H^0(X, \Omega_X)$. This follows from an extension of Raynaud's theory, which identifies $\mathrm{Extrig}(J^0, \mathbb{G}_m)$ with $\mathbb{H}^1(\mathcal{O}_X^* \xrightarrow{d\log} \Omega_X)$ [Co2, §3].

Since the Hecke operators of $X_1(Np)$ give endomorphisms of J and J^0, by Néron's theory, they stabilize the lattice $H^1(J)$ in the de Rham cohomology of $J_1(Np)$ over $\mathbb{Q}_p(\zeta_p)$. Hence they stabilize the lattice $H^1(X)$ in the de Rham cohomology of $X_1(Np)$. We may therefore define $H^1(G) = \varepsilon_m H^1(X) = \varepsilon_m H^1(J)$ and $\omega_G = \varepsilon_m H^0(X, \Omega^1) = \varepsilon_m \omega_J$ as in (15.2), and have an exact sequence of $\mathbb{Z}_p[\zeta_p] \otimes H_m$-modules $0 \to \omega_G \to H^1(G) \to \mathrm{Lie}({}^tG) \to 0$ as in (15.3). The question of a crystalline structure on $H^1(G)$ is a bit more subtle, as J is not an abelian scheme. But, as in the proof of Proposition 12.9, we have an isogeny of abelian varieties $J_1(Np) \to A \times B$ over $\mathbb{Q}$ which has degree prime to p. This isogeny induces a decomposition $H^1(J) = H^1(A) \oplus H^1(B)$ in cohomology over $\mathbb{Z}_p[\zeta_p]$, defined using rigidified extensions and the Néron models of A and B. But A is an abelian scheme over $\mathbb{Z}_p[\zeta_p]$ with special fibre $\bar{A}$; since the ideal $(1 - \zeta_p)$ has divided powers we again have a canonical isomorphism of $\mathbb{Z}_p[\zeta_p]$-modules $H^1(A) = D(\bar{A}) \otimes \mathbb{Z}_p[\zeta_p]$. Since $H^1(A)$ is stable under H_p and $\varepsilon_m H^1(J) = \varepsilon_m H^1(A)$, we obtain an isomorphism of $\mathbb{Z}_p[\zeta_p] \otimes H_m$-modules (if the Hecke actions are compatible...)

$$H^1(G) = D(\bar{G}) \otimes \mathbb{Z}_p[\zeta_p] = U \oplus P \tag{15.7}$$

as in (15.4), and hence an action of Fr_p on $H^1(G)$ commuting with H_m. We summarize the situation in the following Proposition.

PROPOSITION 15.8. *When $k = 2$ let $R = \mathbb{Z}_p$; when $3 \leqslant k \leqslant p$ let $R = \mathbb{Z}_p[\zeta_p]$. There is an exact sequence of $R \otimes H_m$ modules*

$$\begin{tikzcd}
0 \arrow[r] & \omega_G \arrow[r] \arrow[d, equal] & H^1(G) \arrow[r] \arrow[d, equal] & \mathrm{Lie}({}^tG) \arrow[r] & 0 \\
& \varepsilon_m H^0(X, \Omega_X) & \varepsilon_m H^1(X). & &
\end{tikzcd}$$

as well as an action of Frobenius Fr_p *on $H^1(G)$ which commutes with H_m. In the decomposition*

$$H^1(G) = U \oplus P$$

into unit and p-root eigenspaces for Fr_p *the subspace U is complementary to ω_G, and ω_G projects isomorphically onto P. The $R \otimes H_m$-modules $U =$* $\mathrm{Hom}(T_p\bar{G}, R)$, *$P = R \otimes T_p{}^t\bar{G}$, and ω_G are all free of rank* 1, *and $H^1(G)$ is free of rank* 2.

Proposition 15.8 actually holds when R is a complete, local, flat extension of $\mathbb{Z}_p$ (or $\mathbb{Z}_p[\zeta_p]$), defining $H^1(G/R) = H^1(G) \otimes R$, etc. We write U_R, P_R and $\omega_{G/R}$ for the corresponding $R \otimes H_m$ submodules. As in the previous section, we now specialize to the case where R is the completion of the maximal unramified extension of $\mathbb{Z}_p$ (when $k = 2$) or $\mathbb{Z}_p[\zeta_p]$ (when $3 \leqslant k \leqslant p$). The choice of

$$\beta \in R \otimes T_p{}^t\bar{G} \simeq P_R \simeq \omega_{G/R}$$

determines an element $p_\beta \in P_R$ as well as an invariant differential $\omega_\beta = u_\beta + p_\beta$ on G over R with this component. The choice of

$$\alpha \in R \otimes T_p\bar{G} = \mathrm{Hom}(U_R, R)$$

gives an R-linear map $\varphi_\alpha\colon U_R \to R$. Dwork's interpretation of the logarithm of the Serre-Tate invariant is the following (cf. [Dw] for elliptic curves, and [K5] in general).

PROPOSITION 15.9. *For all $\alpha \in R \otimes T_p\bar{G}$ and $\beta \in R \otimes T_p{}^t\bar{G}$ we have*

$$\varphi_\alpha(u_\beta) = \log q'_G(\alpha, \beta) \quad \textit{in } \pi R.$$

Note. When $H_m = \mathbb{Z}_p$ and α is a *basis* of the free R-module $R \otimes T_p\bar{G}$, there is a unique element $u_\alpha \in U_R$ with $\varphi_\alpha(u_\alpha) = 1$. Proposition 15.9 takes the more attractive form

$$\omega_\beta = \log q_G(\alpha, \beta)\cdot u_\alpha + p_\beta. \tag{15.10}$$

In this sense, the image of q_G in $1 + \pi R$ (which reflects the splitting in the sequence $0 \to G^0 \to G \to G^e \to 0$ of p-divisible groups over R) is reflected in the divisibility of the unit eigencomponents of invariant differentials on G over R.

Proof. Dwork's formula actually holds over the formal coordinate ring $\mathscr{R}$ of the formal Lie group $\hat{\mathscr{M}} = \mathrm{Hom}_{\mathbb{Z}_p}(T_p\bar{G} \otimes T_p{}^t\bar{G}, \hat{G}_m)$ of multiplicative type over the Witt vectors W of the residue field of R. The ring $\mathscr{R}$ is the parameter space of the universal formal deformation $\mathscr{G}$ of $\bar{G}$, by the theory of Serre and Tate. If $\{\alpha_i\}$ is a $\mathbb{Z}_p$-basis of $T_p\bar{G}$ and $\{\beta_j\}$ a $\mathbb{Z}_p$-basis of $T_p{}^tG$, then the elements $q_{ij} = \mathbf{q}(\alpha_i, \beta_j)$ are 1-units in $\mathscr{R}$. (Here $\mathbf{q}$ is the pairing associated to the deformation $\mathscr{G}$ over $\mathscr{R}$.) We have an isomorphism $\mathscr{R} \simeq W[[q_{ij} - 1]]$.

The sequence of F-crystals $0 \to U_{\mathscr{R}} \to H^1(\mathscr{G}/\mathscr{R}) \to P_{\mathscr{R}} \to 0$ is split over a larger ring $\tilde{\mathscr{R}}$—the subring of $W \otimes \mathbb{Q}_p[[q_{ij} - 1]]$ consisting of series which converge in the open unit disc (i.e.,—whenever $q_{ij} - 1$ is the maximal ideal of $\mathbb{C}_p$). We have the general formula:

$$\varphi_\alpha(u_\beta) = \log \mathbf{q}(\alpha, \beta) \quad \text{in } \tilde{\mathscr{R}}. \tag{15.11}$$

Since $\tilde{\mathscr{R}}$ is contained in the ring of divided powers of $\mathscr{R}$, the specialization map $\mathscr{R} \to R$ induced by G gives a map $\tilde{\mathscr{R}} \to R$ and Proposition 14.12 follows.

To prove (15.11) we follow Katz [K5]. For each bilinear form $\ell\colon T_p\bar{G} \otimes T_p{}^t\bar{G} \to \mathbb{Z}_p$ there is a unique continuous derivation $D(\ell)$ of $\mathscr{R}$ which is W-linear. By [K5, 3.2] we have $D(\ell)\mathbf{q}(\alpha, \beta) = \ell(\alpha, \beta)\cdot\mathbf{q}(\alpha, \beta)$, which we may write as

$$D(\ell)\log \mathbf{q}(\alpha, \beta) = \ell(\alpha, \beta). \tag{15.12}$$

Let ∇ be the Gauss-Manin connection on $H^1(\mathscr{G}/\mathscr{R})$. For $\beta \in T_p{}^t\bar{G}$, the element p_β in

$H^1(\mathscr{G}/\tilde{\mathscr{R}})$ is characterized by: $\omega_\beta - p_\beta \in U_{\tilde{R}}$, $\nabla p_\beta = 0$. Let $\ell_\beta \in \mathrm{Hom}(T_p\bar{G}, \mathbb{Z}_p)$ be defined by $\ell_\beta(\alpha) = \ell(\alpha, \beta)$ and let $\mathrm{Fix}(\ell_\beta)$ be the corresponding element in U_R. By [K, Thm. 4.3.2] we have

$$(15.13) \qquad \begin{cases} \nabla(D(\ell))\omega_\beta = \mathrm{Fix}(\ell_\beta) \\ \nabla(D(\ell))\mathrm{Fix}(\alpha^\vee) = 0 \quad \text{all}\,\alpha^\vee \text{ in}\,\mathrm{Hom}(T_p\bar{G}, \mathbb{Z}_p). \end{cases}$$

If $\{\alpha_i\}$ is a $\mathbb{Z}_p$-basis of $T_p\bar{G}$, with dual basis $\{\alpha_i^\vee\}$, the element $\omega_\beta - \Sigma_i \log \mathbf{q}(\alpha_i, \beta)\,\mathrm{Fix}(\alpha_i^\vee)$ is annihilated by ∇, hence equal to p_β. Therefore $u_\beta = \Sigma_i \log \mathbf{q}(\alpha_i, \beta)\cdot\mathrm{Fix}(\alpha_i^\vee)$; formula (15.11) then follows from $\varphi_\alpha(\mathrm{Fix}(\alpha_i^\vee)) = \alpha_i^\vee(\alpha)$.

Associated to $\alpha \in R \otimes T_p\bar{G} = \mathrm{Hom}(U_R, R)$ there is a unique $R \otimes H_m$-linear map $\psi_\alpha\colon U_R \to R \otimes H_m^\vee = \mathrm{Hom}_{\mathbb{Z}_p}(H_m, R)$ defined by $\psi_\alpha(u)(h) = \varphi_\alpha(hu)$. As a corollary of Proposition 15.9, we obtain the formula $\psi_\alpha(u_\beta) = \log q_G(\alpha, \beta)$ in $\pi R \otimes H_m^\vee$. Indeed, viewing $\log q_G(\alpha, \beta)$ in $\mathrm{Hom}(H_m, \pi R)$ we have $\log q_G(\alpha, \beta)(h) = \log q'_G(h\alpha, \beta) = \log q'_G(\alpha, h\beta) = \varphi_\alpha(h \cdot u_\beta) = \psi_\alpha(u_\beta)$. Now fix bases h and α of the free H_m-modules $H_m^\vee$ and $T_p\bar{G}$ of rank 1; then there is a unique element $u_\alpha \in U_R$ such that $\psi_\alpha(u_\alpha) = 1 \otimes h$ in $R \otimes H_m^\vee$. Since u_α is a basis of the free $R \otimes H_m$-module U_R, we may write

$$(15.14) \qquad \omega_\beta = c(\alpha, \beta)\cdot u_\alpha + p_\beta$$

with $c(\alpha, \beta) \in R \otimes H_m$. Applying ψ_α, we find that

$$(15.15) \qquad \log q(\alpha, \beta) = c(\alpha, \beta)\cdot(1 \otimes h)$$

in $R \otimes H_m^\vee$, which is the generalization of 15.10 (where $H_m = \mathbb{Z}_p$ and $h = 1$).

Combining the identities (15.14–15.15) with Proposition 14.7, we obtain our final result of this section.

PROPOSITION 15.16. *Let R be a complete local flat $\mathbb{Z}_p[\zeta_p]$-algebra (or $\mathbb{Z}_p$-algebra, when $k = 2$) and let ω be an element of $\omega_{G/R}$. Write $\omega = u + p$ using the decomposition $H^1(G/R) = U_R \oplus P_R$. Then $u \in \pi^{k'}U_R$. If the sequence of E-vector spaces $0 \to G^0[m] \to G[m] \to G^e[m] \to 0$ is split, then $u \in (\pi^{k'}R \otimes mH_m)U_R$.*

Proof. For R the completion of the maximal unramified extension of $\mathbb{Z}_p[\zeta_p]$, this follows directly from the identities. It is therefore true over $\mathbb{Z}_p[\zeta_p]$, by extension of scalars. Finally, it holds for any extension R of $\mathbb{Z}_p[\zeta_p]$, as $\omega_{G/R} = \omega_G \otimes R$, and similarly for U_R, P_R and $H^1(G/R)$. When $k = 2$, the same argument works over $\mathbb{Z}_p$.

§16. Washnitzer-Monsky classes. We begin by reinterpreting Proposition 15.16 using regular differentials on the scheme X. Let F be a lifting of f to an eigenform of weight 2 for $\Gamma_1(N)$ (when $k = 2$) or for $\Gamma_1(Np)$ (when $3 \leqslant k \leqslant p$) as guaranteed by Proposition 9.3. Let R be the complete local $\mathbb{Z}_p$-algebra generated by the coefficients A_n of F, and the values of ε_F. The regular differential $\omega_F = F(q)dq/q$ then lies in the

ε_m-component of $H^0(X, \Omega_{X/R})$; since $\varepsilon_m H^0(X, \Omega_{X/R}) = \omega_{G/R}$ is a free $R \otimes H_m$ module of rank 1, we may write $\omega_F = \omega | h_F$ where ω is a basis and $h_F \in R \otimes H_m$.

The eigenform F determines a ring homomorphism $\varphi_F \colon R \otimes H_m \to R$ which maps $1 \otimes T_\ell$ to A_ℓ, etc. Since $\omega_F | h = \varphi_F(h) \cdot \omega_F$ for all $h \in R \otimes H_m$, we have

$$h_F \cdot h = h_F \cdot \varphi_F(h) \tag{16.1}$$

in $R \otimes H_m$, for all $h \in R \otimes H_m$. We let U^F and P^F denote the R-submodules of U_R and P_R where the algebra $R \otimes H_m$ acts by the character φ_F.

PROPOSITION 16.2. *Let $f = \Sigma a_n q^n$ be a normalized eigenform for $\Gamma_1(N)$ of weight k satisfying $2 \leqslant k \leqslant p$, and assume $a_p \neq 0$. If $k = p$ assume further that $a_p^2 \neq \varepsilon(p)$. Let F be a lifting of f to an eigenform of weight 2 for $\Gamma_1(N)$ or $\Gamma_1(Np)$ with coefficients in R, and let $\omega_F = F(q)dq/q$ be the associated regular differential on X over R.*

Then $\omega_F = u_F + p_F$ in $H^1(X/R)$, where $u_F \in U^F \subset U_R \subset \varepsilon_m H^1(X/R)$ and $p_F \in P^F \subset P_R \subset \varepsilon_m H^1(X/R)$. The unit eigencomponent u_F lies in $\pi^{k'} U^F$. If the representation ρ_f is diagonalizable when restricted to $\mathscr{G}_p = \mathrm{Gal}(\bar{\mathbb{Q}}_p/\mathbb{Q}_p)$, then u_F lies in $m_R \cdot \pi^{k'} U^F$, where m_R is the maximal ideal of R.

Proof. Recall that ω is a basis for the free $R \otimes H_m$-module $\varepsilon_m H^0(X, \Omega_{X/R}) = \omega_{G/R} = U_R \oplus P_R$. Write $\omega = u + p$ with $u \in U_R$ and $p \in P_R$. Then $u_F = u | h_F$ lies in U^F by (16.1); similarly $p_F = p | h_F$ lies in P^F and $\omega_F = u_F + p_F$.

Since $u \in \pi^{k'} U_R$ by Proposition 15.16, we clearly have $u_F \in \pi^{k'} U^F$. If ρ_f is completely reducible when restricted to $\mathscr{G}_p$, the sequence $0 \to G^0[m] \to G[m] \to G^e[m] \to 0$ of E-vector space schemes is split, and $u = (\pi^{k'} \otimes \alpha) u_0$ with $\alpha \in mH_m$ and $u_0 \in U_R$. Since φ_F maps the ideal $\pi^{k'} R \otimes mH_m$ into the ideal $\pi^{k'} \cdot m_R$, and $u_0 | h_F$ is contained in U^F, this shows that $u_F \in m_R \pi^{k'} U^F$.

For the rest of this section, we will assume that ρ_f is diagonalizable when restricted to $\mathscr{G}_p$, so $u_F \in m_R \pi^{k'} U^F$. We now unify the cases $k = 2$ and $3 \leqslant k \leqslant p$ by passing to the eigenform F' on $\Gamma_1(Np)$. This is an old form, defined by (9.10) when $k = 2$, and a newform defined by the equation $F | w_\zeta = c_\zeta \cdot F'$ when $k \neq 2$. The differential $\pi^{k'} \cdot \omega_{F'}$ is regular on the model X of $X_1(Np)$ over R, by Proposition 9.13, part 3).

COROLLARY 16.3. *We have $\pi^{k'} \cdot \omega_{F'} = u_{F'} + p_{F'}$ in $H^1(X/R)$, where $u_{F'}$ and $p_{F'}$ are in the F'-eigenspace of the unit-root and p-root eigenspaces for* Fr_p. *Furthermore, $u_{F'} = \pi^{k'} \cdot \alpha \cdot u'$ with $u' \in U^{F'}$ and $\alpha \in m_R$.*

Proof. When $k \neq 2$ this follows from the identity $\omega_F | w_\zeta = c_\zeta \cdot \omega_{F'}$ proved in Proposition 6.14 and Proposition 16.2. Indeed, c_ζ has the same p-adic valuation as $\pi^{k'}$, and w_ζ preserves the eigenspaces for Fr_p acting on $H^1(A/R) \subset H^1(X/R)$.

When $k = 2$ the matter is a bit more subtle, as the class $\pi^{k'} \omega_{F'}$ lies in the summand $H^1(B/R)$ of $H^1(X/R)$, where we have not yet defined a crystalline structure. But B is isogenous, by an isogeny of degree prime to p, to the product $J_1(N) \times J_1(N)' \times K_1$, where K_1 has multiplicative reduction at p. Since $\pi^{k'} \omega_{F'}$ corresponds to an invariant differential on $J_1(N) \times J_1(N)'$, it lies in $H^1(J_1(N)/R) \oplus H^1(J_1(N)'/R)$ which has an action of Fr_p. The corollary then follows from Proposition 16.2, using the formula

$$p \cdot \omega_{F'} = p \cdot \pi^*(\omega_F) - u\pi'^*(\omega_F). \tag{16.4}$$

In (16.4) u is the unit root of $x^2 - A_p x + p\varepsilon_N(p) = 0$ and $\pi, \pi'\colon X_0(Np) \to X_0(N)$ are the usual coverings $\pi(E, \alpha, \beta) = (E, \alpha)$, $\pi'(E, \alpha, \beta) = (E', \alpha')$ with $E' = E/\beta(\mu_p)$ and α' the induced injection of μ_N.

To exploit the extra divisibility in $u_{F'}$, we will consider a map η from $H^1(X/R)$ to the Washnitzer-Monsky cohomology of the affinoid subdomains V and V' consisting of points reducing to $I - \Sigma$ and $I' - \Sigma$, respectively (modulo m_R). The scheme X is obtained by glueing the two "wide open spaces" $\bar{V}$ and $\bar{V}'$ with these underlying affinoids along their intersection: the annuli reducing to points in Σ [Col]. Figure 1 illustrates the analogy of the map η with the Mayer-Vietoris sequence for computing the de Rham cohomology of a Riemann surface.

Let V_n be the reduction of V (mod m_R^n), which is an affine scheme over R/m_R^n, and let $\mathscr{A}(V) = \varprojlim_n H^0(V_n, \mathcal{O}_n)$. This is an R-algebra, with $\mathscr{A}(V) \otimes R/m_R$ isomorphic to the coordinate ring of the affine curve $I - \Sigma$ over E. Let K be the quotient field of R; then $\mathscr{A}(V) \otimes_R K$ is the algebra of rigid analytic functions $\mathscr{F}$ on V over K. An analytic differential $\mathscr{F} \cdot d\mathscr{G}$ on V over K is said to be over-convergent if it extends to some neighborhood of V in $\bar{V}$. The Washnitzer-Monsky cohomology group $H^1(V/K)$ is defined as the quotient of the space of over-convergent differentials by the exact differentials [WM]. This K-vector space is finite dimensional and has a

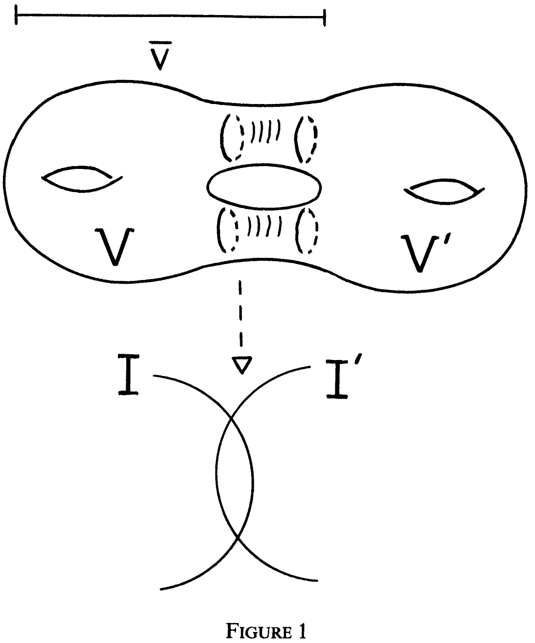

FIGURE 1

K-linear action of Fr_p: it depends functorially only on the reduced curve $I - \Sigma$. As such, it has an action of the Hecke algebra H_p.

Define the K-linear mapping

$$\eta\colon H^1(X/K) \to H^1(V/K) \tag{16.6}$$

on a class c by first representing c by a differential v of the second kind on X (modulo exact differentials) whose poles reduce to points in $I' - \Sigma$. Then v is analytic on $\bar{V}$, and we may define $\eta(c) = v|_V$. I will assume that the map η commutes with the action of H_p and Fr_p on each factor. This is undoubtedly true, but I could not find it in the literature; it would follow from a general theory of correspondences in p-adic cohomology. Hence η induces a map of the F'-eigenspaces: $H^1(X/K)^{F'} \to H^1(V/K)^{F'}$ which we may apply to the differential $\omega_{F'}$.

PROPOSITION 16.7. *We have $\eta(\omega_{F'}) = \alpha \cdot \eta(u')$ in $H^1(V/K)^{F'}$, where $u' \in U^{F'}$ is in the unit eigenspace for Fr_p (over R) and $\alpha \in m_R$.*

Proof. Since $\omega_{F'} = \alpha \cdot u' + p'$ with $p' = \pi^{-k'} p_{F'}$ by Corollary 16.3, it suffices to show that $P^{F'}$ is in the kernel of η. To do this, it suffices to show that Fr_p is a unit on the eigenspace $H^1(V/K)^{F'}$. But $U_p \equiv \mathrm{Ver}_p$ on I, so Ver_p acts as $A'_p = p\varepsilon_N(p)/A_p$ on $H^1(V/K)^{F'}$. Since A'_p has the same p-adic valuation as p, $\mathrm{Fr}_p = p/\mathrm{Ver}_p$ acts as the unit $A_p/\varepsilon_N(p)$.

Let v' be a differential on X over R which represents the class of u' in U^F and has poles reducing to points in $I' - \Sigma$. The following result completes the proof of Theorem 13.11.

PROPOSITION 16.8. 1) *The expansions $\omega_{F'} = \Sigma A'_n q^n dq/q$ and $v' = \Sigma C_n q^n dq/q$ at ∞ are both integral and formally exact over R: $A'_n \equiv C_n \equiv 0 \,(\mathrm{mod}\ nR)$ for all $n \geqslant 1$.*

2) *If ρ_f is completely reducible when restricted to $\mathscr{G}_p$, then the reduction of the q-expansion $\Sigma(A'_n/n)q^n$ $(\mathrm{mod}\ m_R)$ is the Fourier expansion of a modular for g for $\Gamma_1(N)$ over E. The form g lies in $\tilde{M}_{k'}$ and satisfies: $\theta g = \theta^{k'} f$, $g|U_p = \lambda g$ with $\lambda = \varepsilon(p))/a_p \neq 0$. It has filtration k' (or p, if $k' = 1$) and is a companion to the eigenform f.*

Proof. 1) By part 2) of Proposition 9.13, the expansion $\Sigma A'_n q^n dq/q$ is formally exact. The expansion of v' at ∞ is integral, as $v' \in H^1(X/R)$ and ∞ is a smooth section of X over R. Since v' is an eigenvector for Fr_p with eigenvalue $\lambda = A_p/\varepsilon_N(p)$ we have

$$p\Sigma C_n q^{np} dq/q = \lambda \cdot \Sigma C_n q^n dq/q + d(\Sigma D_n q^n)$$

with $D_n \in R$ for all $n \geqslant 1$. Hence $\lambda C_{np} + npD_{np} = pC_n$ for all $n \geqslant 1$. Since λ is a unit, an induction on the power of p dividing n shows that $C_n \equiv 0$ (mod nR). For more on the expansion of classes in the unit root eigenspace, see [K4].

2) When ρ_f is reducible when restricted to $\mathscr{G}_p$, by Proposition 16.7 we have $\omega_{F'} = \alpha \cdot v' + d\mathscr{F}$ as an equality of overconvergent differentials on V, where $\alpha \in m_R$

and $\mathscr{F}$ is a rigid analytic function on V. Take the expansions at ∞ to obtain

$$\Sigma A'_n q^n dq/q = \alpha \cdot \Sigma C_n q^n dq/q + \Sigma n B_n q^n dq/q$$

where $\mathscr{F} = \Sigma B_n q^n$. Thus $A'_n = \alpha C_n + nB_n$; by part 1) we have $B_n \in R$ for all $n \geqslant 1$ and hence $\mathscr{F}$ is a rigid function in the R-algebra $\mathscr{A}(V)$. Since $\alpha \in m_R$ we have the congruence: $A'_n/n \equiv B_n \pmod{m_R}$ so $\Sigma(A'_n/n)q^n \equiv \mathscr{F}(q)$ (modulo $m_R\mathscr{A}(V)$). But the elements in $\mathscr{A}(V)/m_R\mathscr{A}(V)$ are, by definition, in the affine ring of the curve $I - \Sigma$ over E. Hence the reduction of the expansion $\Sigma(A'_n/n)q^n$ is the Fourier expansion of a modular form g for $\Gamma_1(N)$ over E, by Proposition 5.5!

Since $\Sigma A'_n q^n$ is an eigenvector for U_p with eigenvalue A'_p, and $(\theta H)|U_p = p \cdot \theta(H|U_p)$ for any expansion with p-adic coefficients, the formal expansion $H = \Sigma(A'_n/n)q^n$ is an eigenvector for U_p with eigenvalue A'_p/p. Indeed, $\theta(H|U_p - (A'_p/p)H) = 0$ and the kernel of θ is characteristic zero is the constants. Since H has zero constant term $H|U_p - (A'_p/p)H = 0$. Hence $g(q)$ is an eigenvector for U_p with eigenvalue $\lambda = \varepsilon(p)/a_p \equiv A'_p/p \pmod{m_R}$. Since $a_p^2 \neq \varepsilon(p)$ when $k = p$, $g \neq f$. Similarly, one can show that g is an eigenvector for T_ℓ acting on $\tilde{M}_{k'}$, for all $\ell \nmid N$.

We have $\theta g = \theta^{k'} f \equiv F' \pmod{m_R}$ by part 1) of Proposition 9.13. Write $g = \Sigma g_\alpha$ using the direct sum decomposition $\tilde{M} = \oplus \tilde{M}_\alpha$. Since θg lies in $\tilde{M}_{k'+p+1}$, we have $g_\alpha = h_\alpha | V_p$ for all $\alpha \not\equiv k' \pmod{p-1}$. Apply U_p and use the identity $V_p U_p = 1$ to get $\lambda g = \Sigma \lambda g_\alpha = g_{k'}|U_p + \Sigma_{\alpha \neq k'} h_\alpha$. Hence $h_\alpha = \lambda \cdot h_\alpha | V_p$ for all $\alpha \not\equiv k'$. Taking the q-expansion, we find that $h_\alpha = 0$, unless $\lambda = 1$, $\alpha \equiv 0$ and $h_\alpha =$ constant. But even in the latter case, we must have $h_\alpha = 0$, as h_α is an eigenvector for T_ℓ with eigenvalue $a_\ell \cdot \ell^{k'-1}$. If $h_\alpha \neq 0$, we have $a_\ell \cdot \ell^{k'-1} = 1 + \varepsilon(\ell)\ell^{k'-1}$ for all $\ell \nmid Np$, which implies that $\rho_f \simeq \varepsilon \oplus \chi^{k-1}$ is reducible (a contradiction). So $g = g_{k'}$ lies in $\tilde{M}_{k'}$.

Since $g|U_p = \lambda g$ with $\lambda \neq 0$, the series $g(q)$ has filtration m satisfying $2 \leqslant m \leqslant p + 1$ by Proposition 4.12. Since θg has filtration $k' + p + 1$, we must have $m < p + 1$ and hence $m = k'$ (or $m = p$, if $k' = 1$). Therefore g is the desired companion to f.

§17. Examples. We now give some examples of cuspidal eigenforms (mod p) of weight $k \leqslant p + 1$ on $\Gamma_1(N)$, discuss thier liftings to forms of weight 2 on $\Gamma_1(Np)$ and their Galois representations, and describe the search for a companion form.

The cusp form

$$\Delta = q \prod_{n \geqslant 1} (1 - q^n)^{24} = \sum_{n \geqslant 1} \tau(n) q^n \tag{17.1}$$

has weight 12 for $\Gamma_1(1) = SL_2(\mathbb{Z})$ and is defined over $\mathbb{Z}$. It gives an eigenform (mod p) for all primes p, but we will assume that $p \geqslant 11$ so as to have the inequality $k \leqslant p + 1$.

When $p = 691$, Δ is congruent to the Eisenstein series of weight 12 for $SL_2(\mathbb{Z})$: $\tau(n) \equiv \sigma_{11}(n) \pmod{691}$. This implies that the Galois representation ρ_Δ is reducible, and isomorphic to $1 \oplus \omega^{11}$. When $p = 23$, Δ is equal to its own companion form:

$\tau(n) \equiv n^{11}\tau(n) \equiv \left(\frac{n}{23}\right)\tau(n) \pmod{23}$ for all $(n, 23) = 1$. In this case, the representation ρ_Δ has image the symmetric group on 3 letters in $GL_2(\mathbb{Z}/23\mathbb{Z})$, and describes the maximal abelian unramified extension of $\mathbb{Q}(\sqrt{-23})$. In all other case (i.e., $p \geqslant 11$, $p \neq 23$ or 691) the representation ρ_Δ: $Gal(\bar{\mathbb{Q}}/\mathbb{Q}) \to GL_2(\mathbb{Z}/p\mathbb{Z})$ has image the subgroup of invertible matrices A with $\det A \in (\mathbb{Z}/p\mathbb{Z})^{*11}$ [S4, §3]. In particular, ρ_Δ determines a Galois extension of $\mathbb{Q}(\mu_p)$ which is unramified outside p and has Galois group $\simeq SL_2(\mathbb{Z}/p\mathbb{Z})$.

For $p = 11$ we have $k = p + 1$. By the results of §9 there is a lifting of $f = \Delta$ to a form F of weight 2 and trivial character on $\Gamma_1(11)$. The unique such form has q-expansion

$$F = q\prod_{n\geqslant 1} (1 - q^n)^2(1 - q^{11n})^2.$$

The representation ρ_Δ occurs on the 11-division points of the elliptic curve $J_0(11)$ (or equivalently, on the 11-division points of the 5-isogenous curve $J_1(11)$). Its restriction to the inertia group at 11 is "très ramifié" in the sense of Serre [S8, pg. 186].

When $p = 13$ we have $k = p - 1$. By the results of §5 there is a lifting of $f = \Delta$ to a form F of weight 2 and character ω^{10} on $\Gamma_1(13)$. The unique such form has q-expansion in the subring $\mathbb{Z}[\mu_6]$ of $\mathbb{Z}_{13}$. If α is the unique 6^{th} root of 1 in $\mathbb{Z}_{13}^*$ which satisfies $\alpha \equiv 4 \pmod{13}$ then the Fourier expansion of F begins

$$F = q + (-2 + \alpha)q^2 + (-2\alpha)q^3 + (1 - \alpha)q^4 + (2\alpha - 1)q^5 + (2 + 2\alpha)q^6 + \cdots.$$

The form F' is the complex conjugate of F; it has weight 2 and character ω^2. The Fourier expansion of F' is obtained by replacing α by $\bar{\alpha} = 1 - \alpha$ in the above

$$F' = q + (-1 - \alpha)q^2 + (-2 + 2\alpha)q^3 + \alpha q^4 + (1 - 2\alpha)q^5 + (4 - 2\alpha)q^6 + \cdots.$$

We have $A_{13} = -4 + 3\alpha \equiv 8 \pmod{13}$ and $A'_{13} \equiv -1 - 3\alpha \equiv 0 \pmod{13}$. The form F is congruent to the eigenform $f = \Delta_{12} \pmod{13}$, and F' is congruent to the eigenform $f' = \theta^2 f = \Delta_{16} \pmod{13}$, where Δ_{16} is the unique normalized cusp form of weight 16 and level 1. The differential $\nu_{f'}$ is meromorphic on the Igusa curve $I = I_1(1)$ over $\mathbb{Z}/13\mathbb{Z}$; it has a pole of order 2 at the unique supersingular point and is exact: $\nu_{f'} = dg$. Here the function $g = \theta f = \Delta_{26} \pmod{13}$ has a pole of order 13 at the unique supersingular point. The representation $\rho_\Delta = \rho_f$ occurs in the subspace of 13-division points of $J_1(13)$ where the Galois group $(\mathbb{Z}/13\mathbb{Z})^*/\langle \pm 1\rangle$ of $X_1(13)$ over $X_0(13)$ acts by ω^2. As there is no companion form, ρ_f gives an $SL_2(\mathbb{Z}/13\mathbb{Z})$-extension of $\mathbb{Q}(\mu_{13})$, which is wildly ramified at 13.

A computer search by Elkies, extended by Atkin, showed that Δ does *not* have a companion form for $p < 3{,}500$ with $p \geqslant 11$ and $p \neq 23$ or 691. Hence the representation ρ_Δ gives an $SL_2(\mathbb{Z}/p)$-extension of $\mathbb{Q}(\mu_p)$ which is ramified at p in all these cases.

The same search by Elkies and Atkin *did* discover companion forms for the eigenforms of weight 16, 18, 20, and 26 for $SL_2(\mathbb{Z})$ with integral Fourier coefficients.

Let E_4 and E_6 be the normalized Eisenstein series

$$E_4 = 1 + 240\Sigma\sigma_3(n)q^n$$

$$E_6 = 1 - 504\Sigma\sigma_5(n)q^n,$$

and define the cusp forms

$$\Delta_{16} = E_4\Delta$$

$$\Delta_{18} = E_6\Delta$$

$$\Delta_{20} = E_4^2\Delta$$

$$\Delta_{22} = E_4E_6\Delta$$

$$\Delta_{26} = E_4^2E_6\Delta.$$

The following table lists the primes $p < 3{,}500$ where the image of ρ_f contains $SL_2(\mathbb{Z}/p\mathbb{Z})$ and f has a companion form.

f	companion $p < 3{,}500$
$\Delta = \Delta_{12}$	none
Δ_{16}	397
Δ_{18}	271
Δ_{20}	139, 379
Δ_{22}	none
Δ_{26}	107.

As an example, let us describe the companion form g of weight 82 for $f \equiv \Delta_{26}$ (mod 107), which was discovered by Elkies. (He also checked that $p = 107$ is the smallest prime where there are a pair of companions (f, g) of level $N = 1$.) We have

$$g \equiv E_4E_6\Delta(E_4^{15} - 15E_4^{12}\Delta - 35E_4^9\Delta^2 + 36E_4^6\Delta^3 - 18E_4^3\Delta^4 + 15\Delta^5)$$

and $a_n \equiv n^{25}b_n \pmod{107}$ for $(n, 107) = 1$. In this case, $a_{107} \equiv b_{107} \equiv -1 \pmod{107}$.

The Fourier expansions of f and g begin:

$$f \equiv q - 48q^2 + 6q^3 - 31q^4 + 45q^5 + 33q^6 + 10q^7 + 41q^8 + 38q^9 - 20q^{10}$$
$$- 38q^{11} + 28q^{12} - 48q^{13} + \cdots$$

$$g \equiv q - 20q^2 - 3q^3 + 34q^4 + 12q^5 - 47q^6 - 49q^7 + 6q^8 - 44q^9 - 26q^{10}$$
$$- 43q^{11} + 5q^{12} + 50q^{13} + \cdots$$

This is sufficient to check the identity $\theta f \equiv \theta^{26} g$, which holds in the 10-dimensional space of cusp forms of weight 134.

Next, consider the case when $f = \Sigma a_n q^n$ is the reduction (mod p) of a newform $F = \Sigma A_n q^n$ of weight 2 and trivial character on $\Gamma_1(N)$. Assume further that the Fourier coefficients A_n of F are all rational integers; then F corresponds to an elliptic curve E of conductor N over $\mathbb{Q}$ which occurs in the Jacobian of the curve $X_0(N)$ [Sh 2, Ch. 7] and ρ_f occurs on the p-torsion of E. We assume ρ_f is irreducible as usual.

The condition that $a_p \neq 0$ implies that E has ordinary reduction at p [S3, §4]. Let j_E be the modular invariant of E in $\mathbb{Z}_p$, and let j_0 be the "canonical lifting" of the reduction of j_E (mod p). Then j_0 is the modular invariant of the unique elliptic curve E_0 which satisfies $E_0 \equiv E$ (mod p) and $\mathrm{End}_{\mathbb{Z}_q}(E_0) \equiv \mathrm{End}_{\mathbb{Z}/p\mathbb{Z}}(E)$. A simple argument, along the same lines as §14–15, shows that the restriction of ρ_f to a decomposition group at p is diagonalizable if and only if $j_E = j_0$ (mod p^2) for p odd, and $j_E \equiv j_0$ (mod 8) for $p = 2$. Similarly, the local action on p^n-torsion is diagonalizable if and only if $j_E \equiv j_0$ (mod $2 \cdot p^{n+1}$). In these cases, we will have a companion form of weight $k' = p - 1$ on $\Gamma_0(N)$ over $\mathbb{Z}/p\mathbb{Z}$, whose Fourier coefficients satisfy $nb_n = a_n$ for $(n, p) = 1$.

If $p = 2$ we have $j_E \equiv 1$ (mod 2) and $j_0 = -3^3 5^3$. Hence ρ_f is diagonalizable (= trivial) on the decomposition group at 2 if and only if $j_E = 1$ (mod 8); when $j_E = 1$ (mod 4) one finds that ρ_f is unramified at 2. Since $j_E - 2^6 \cdot 3^3 = c_6^2/\Delta$ and $c_6^2 \equiv 1$ (mod 8), we have:

$$\rho_f \text{ is unramified at } 2 \Leftrightarrow \Delta \equiv 1 \pmod 4$$

$$\rho_f \text{ is diagonalizable at } 2 \Leftrightarrow \Delta \equiv 1 \pmod 8.$$

Since we are in the case when $k = p$ and $a_p^2 = \varepsilon(p)$, Theorem 13.10 does *not* apply. But there is a result analogous to Corollary 13.11: when $\Delta \equiv 1$ (mod 4) there is an eigenform $h = \Sigma c_n q^n$ of weight 1 on $\Gamma_0(N)$ over $\mathbb{Z}/2\mathbb{Z}$ with $c_n = a_n$ for n odd and $c_2 = a_2 + \varepsilon(2)/a_2 = 2a_2 = 0$. Indeed, in this case ρ_f defines a $GL_2(\mathbb{Z}/2\mathbb{Z}) = S_3$ extension of $\mathbb{Q}$ which is unramified at 2, so yields a form of weight 1 on $\Gamma_1(N)$ with the desired Fourier coefficients (mod 2) [S8, §5.1]. The first examples of curves E with $\Delta \equiv 1$ (mod 4), ordinary reduction at 2, and irreducible representation on 2-torsion occur at levels $N = 83, 139$ where $\Delta = -83, -139$ respectively. A case when $\Delta \equiv 1$ (mod 8) and the above hypotheses hold occurs at level $N = 431$, where $\Delta = -431$.

When $p = 3$ we have $j_E \equiv 1, -1$ (mod 3) and $j_0 = -2^{15}, 2^6 5^3$ respectively. Hence ρ_f is diagonalizable on the decomposition group at 3 if and only if $j_E \equiv \pm 1$ (mod 9). Since $j_E = c_4^3/\Delta$ and $c_4^3 \equiv \pm 1$ (mod 9), we have:

$$\rho_f \text{ is diagonalizable at } 3 \Leftrightarrow \Delta \equiv \pm 1 \pmod 9.$$

In this case, Theorem 13.10 applies and there is a companion from $g = \Sigma b_n q^n$ of weight $k' = 2$ on $\Gamma_0(N)$ over $\mathbb{Z}/3\mathbb{Z}$ whose Fourier coefficients satisfy $b_n = \left(\frac{n}{3}\right) a_n$ for

$(n, 3) = 1$, and $b_3 a_3 = 1$. The first case when $\Delta \equiv \pm 1(9)$, E is ordinary at 3, and the representation on 3-torsion is irreducible occurs at level $N = 89$. Here we may take the curve $E = 89C$ with $\Delta = -89$ and associated eigenform.

$$F = q - q^2 - q^3 - q^4 - q^5 + q^6 - 4q^7 + 3q^8 + \cdots.$$

The companion g of f also lifts to an eigenform with integral coefficients:

$$G = q + q^2 + 2q^3 - q^4 - 2q^5 + 2q^6 + 2q^7 - 3q^8 + \cdots.$$

which corresponds to the elliptic curve $E' = 89A$ with $\Delta' = -89^2$.

References

[AS] A. Ash and G. Stevens, *Cohomology of arithmetic groups and congruences between systems of Hecke eigenvalues*, J. Reine Angew. Math. **365**, 192–220 (1986).

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

[ALi] A. O. L. Atkin and W. Li, *Twists of newforms and pseudo eigenvalues of W-operators*, Invent. Math. **48**, 221–244 (1978).

[B] H. Bass, *On the ubiquity of Gorenstein rings*, Math. Z. **82**, 8–28 (1963).

[BM] P. Berthelot and W. Messing, *Théorie de Dieudonné cristalline I.*, Journées de Géométrie Algébrique de Rennes vol. I, Astérisque **63**, 17–37 (1979).

[BBM] P. Berthelot, L. Breen and W. Messing, *Théorie de Dieudonné cristalline II.*, Lecture Notes in Math. **930** (1982).

[Br] L. Breen, *Rapport sur la théorie de Dieudonné*, Journées de Géométrie Algébrique de Rennes, Astérisque **63**, 39–66 (1979).

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

[Co1] R. Coleman, Letters to B. Gross, January 17, 1988, June 17, 1989.

[Co2] ———, *The Universal Vectorial Bi-extension*, To appear, Inv. Math. (1990).

[D1] P. Deligne, *Formes modulaires et représentations ℓ-adiques*, Sém. Bourbaki **355**, Lecture Notes in Math. **179**, 136–172 (1971).

[D2] ———, Letter to J.-P. Serre, May 28, 1974.

[D3] ———, *Courbes elliptiques*, Lecture Notes in Math. **476**, 53–74 (1975).

[DR] P. Deligne and M. Rapoport, *Schémas de modules de courbes elliptiques*, Lecture Notes in Math. **349**, 143–316 (1973).

[DS] P. Deligne and J.-P. Serre, *Formes modulaires de poids 1*, Ann. Sci. Ecole Norm. Sup. (4) **7**, 507–530 (1974). (= J.-P. Serre Oe. 101)

[Dw] B. Dwork, *p-adic cycles*, Inst. Hautes Études Sci. Publ. Math. **37**, 27–115 (1969).

[E] M. Eichler, *Quaternäire quadratische Formen und die Riemannsche Vermutung für die Kongruenzzetafunktion*, Arch. Math. **5**, 355–366 (1954)

[F1] J.-M. Fontaine, *Sur la construction du module de Dieudonné d'un groupe formel*, C.R. Acad. Sci. Paris, Sér. A–B 280, 1273–1276 (1975)

[F2] ———, *Groupes p-divisibles sur les corps locaux*, Astérisque **47–48** (1977)

[F3] ———, Letters to J.-P. Serre, June 25, 1979, July 10, 1979.

[G1] A. Grothendieck, *Groupes de Barsotti-Tate et cristaux*, Actes Congrès International Mathématiciens (Nice) 1970 Tome 1, 431–436.

[G2] ———, *Groupes de Barsotti-Tate et Cristaux de Dieudonné*, Sém. Math. Sup. **45**, Presses Univ. Montreal (1974).

[H] R. Hartshorne, *Algebraic Geometry*, Graduate Texts in Math. **52** (1977).

[J] N. Jochnowitz, *The local components of the Hecke algebra mod ℓ*, Trans. Amer. Math. Soc. **270**, 253–267 (1982).

[K1] N. Katz, *Travaux de Dwork*, Sém. Bourbaki 1971/1972, exp. **409**, Lecture Notes in Math. **317**, 167–200 (1973).

[K2] ———, *p-adic properties of modular schemes and modular forms*, Lecture Notes in Math. **350**, 69–170 (1973).

[K3] ———, *A result on modular forms in characteristic p*, Lecture Notes in Math. **601**, 53–61 (1976).

[K4] ———, "Crystalline cohomology, Dieudonné modules, and Jacobi sums," *Automorphic forms, representation theory, and arithmetic*, Tata Inst. Fund. Res. Studies in Math. **10**, 165–246 (1979).

[K5] ———, *Serre-Tate local moduli*, Lecture Notes in Math. **868**, 138–202 (1981).

[KM] N. Katz and B. Mazur, *Arithmetic moduli of elliptic curves*, Annals of Math. Stud. **108**, Princeton University (1985).

[L] S. Lang, *Cyclotomic Fields I*, Graduate Texts in Math. **59** (1978).

[La] R. P. Langlands, *Modular forms and ℓ-adic representations*, Lecture Notes in Math. **349**, 361–500 (1973).

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

[M] B. Mazur, *Modular curves and the Eisenstein ideal*, Inst. Hautes. Études Sci. Publ. Math. **47**, 33–186 (1977).

[MM] B. Mazur and W. Messing, *Universal extensions and one-dimensional crystalline cohomology*, Lecture Notes in Math. **370** (1974).

[MR] B. Mazur and K. Ribet, *Two dimensional representations in the arithmetic of modular curves*, To appear, Astérique (1990).

[MT] B. Mazur and J. Tate, *Points of order 13 on elliptic curves*, Inv. Math. **22**, 41–49 (1973).

[MW] B. Mazur and A. Wiles, *Class fields of abelian extensions of* $\mathbb{Q}$, Inv. Math. **76**, 179–330 (1984).

[Me] W. Messing, *The crystals associated to Barsotti-Tate groups: with applications to abelian schemes*, Lecture Notes in Math. **264** (1972).

[Ra] M., Raynaud, *Spécialisation du foncteur de Picard*, Inst. Hautes Études Sci. Publ. Math. **38**, 27–76 (1970).

[R] K. Ribet, *On modular representations of* Gal($\bar{\mathbb{Q}}/\mathbb{Q}$) *arising from modular forms*, Inv. Math. **100**, 431–476 (1990).

[S1] J.-P., Serre, *Groupes algébriques et corps de classes*, Hermann (1959).

[S2] ———, *Une interprétation des congruences relatives à la fonction τ de Ramanujan*, Seminar Delange-Pisot-Poitu: 1967/68, Theórie des Nombres, Fasc. 1, Exp. 14, Secretariat mathématique **14** (1967) (= Oe. 80).

[S3] ———, *Propriétés galoisiennes des points d'ordre fini des courbes elliptiques*, Inv. Math. **15**, 259–331 (1972) (= Oe. 94).

[S4] ———. *Congruences et formes modulaires [d'après H.P.F. Swinnerton-Dyer]*, Sém. Bourbaki **416** Lect. Notes. In Math. **317**, 319–338 (1971) (= Oe. 95).

[S5] ———. *Formes modulaires et fonctions zêta p-adiques*, Lecture Notes in Math. **350**, 191–268 (1973) (= Oe. 97).

[S6] ———, *Valeurs propres des operateurs de Hecke modulo ℓ*, Journées Arith. de Bordeaux, (Conf. Univ. Bordeaux, 1974), Astérisque **24–25**, 109–117 (1975) (= Oe. 104).

[S7] ———, Letter to J.-M. Fontaine, May 27, 1979.

[S8] ———, *Sur les représentations modulaires de degré 2 de* Gal($\bar{\mathbb{Q}}/\mathbb{Q}$), Duke Math. J. **54**, 179–230 (1987).

[S9] ———, *Résumé des Cours de* 1987–1988, Annuaire du Collège de France (1988).

[ST] J.-P. Serre and J. Tate, *Good reduction of abelian varieties*, Ann. of Math. (2) **88**, 492–517 (1968).

[Sh1] G. Shimura, G., *Correspondances modulaires et les fonctions ζ de courbes algébriques*, J. Math. Soc. Japan **10**, 1–28 (1958).

[Sh2] ———. *Introduction to the arithmetic theory of automorphic functions*, Publications of the Mathematical Society of Japan **11**, Princeton Univ. Press (1971).

[Sh3] ———, *An ℓ-adic method in the theory of automorphic forms* (unpublished, 1968).

[Sw] H. P. F. Swinnerton-Dyer, *On ℓ-adic representations and congruences for coefficients of modular forms*, Lecture Notes in Math. **350**, 1–55 (1973).

[T] J. Tate, "p-divisible groups," *Proceedings of a Conference on Local Fields at Driebergen*, Springer-Verlag, 158–184 (1967).

[WM] G. Washnitzer and P. Monsky, *Formal Cohomology I.*, Ann. of Math. (2) **88**, 181–217 (1968).

[W] A. Weil, *Variétés abéliennes et courbes algébriques*, Hermann (1948).

[Wi] A. Wiles, *Modular curves and the class group of* $\mathbb{Q}(\zeta_p)$, Inv. Math. **58**, 1–35 (1980).

Department of Mathematics, Science Center 325, Harvard University, Cambridge, Massachusetts 02138.