de Shalit and Goren *Res Math Sci(2016)3:28*
DOI 10.1186/s40687-016-0075-8

Research in
the Mathematical Sciences

RESEARCH Open Access

CrossMark

A theta operator on Picard modular forms modulo an inert prime

Ehud de Shalit[1*] and Eyal Z. Goren[2]

*Correspondence:
deshalit@math.huji.ac.il
[1] Hebrew University, Jerusalem, Israel
Full list of author information is available at the end of the article

To the memory of Robert Coleman

Contents

Swinnerton-Dyer [38] and Serre [34] introduced a certain differential operator θ on (elliptic) modular forms over $\bar{\mathbb{F}}_p$. In terms of the q-expansion

$$f = \sum_{n=0}^{\infty} a_n q^n \tag{0.1}$$

($a_n \in \bar{\mathbb{F}}_p$) of such a form, θ is given by qd/dq. It lifts, by the same formula, to the space of p-adic modular forms. This suggests a relation with the Tate twist of the mod p Galois representation attached to f, if the latter is a Hecke eigenform.

Over $\mathbb{C}$, this operator has been considered already by Ramanujan, where it fails to preserve modularity "by a multiple of E_2". Maass modified it so that modularity is preserved, sacrificing holomorphicity. Shimura studied Maass' differential operators on more general symmetric domains, as well as their iterations. They have become known as Maass–Shimura operators and play an important role in the theory of automorphic forms [37, chapter III].

At the same time, Serre's p-adic operator has been studied in relation to mod p Galois representations, congruences between modular forms, p-adic families of modular forms and p-adic L-functions. As an example, we cite Coleman's celebrated classicality theorem, asserting that "overconvergent modular forms of small slope are classical" [6]. A key step in Coleman's original proof of that theorem was the observation that, although the p-adic theta operator did not preserve the space of overconvergent modular forms, for any $k \geq 0$, θ^{k+1} mapped overconvergent forms of weight $-k$ to overconvergent forms of weight $k+2$.

Underlying the p-adic theory is Katz' geometric approach to the theta operator, via the Gauss–Manin connection on the de Rham cohomology of the universal elliptic curve [20,21]. Broadly speaking, Katz' starting point is the unit-root splitting of the Hodge filtration in this cohomology over the ordinary locus. It is supposed to replace the Hodge decomposition over $\mathbb{C}$, which can be used to make a *geometric* theory of the C^∞ operators of Maass–Shimura, thereby explaining their arithmetic significance. This approach has been adapted successfully to other Shimura varieties of PEL type, as long as they admit a non-empty ordinary locus in their characteristic p fiber. For unitary Shimura varieties, this has been done by Eischen [9,10], if *p splits* in the quadratic imaginary field [and the signature is (n, n)]. Böcherer and Nagaoka [3] defined theta operators on Siegel modular forms by studying their q-expansions.

The assumption that the ordinary locus is non-empty may nevertheless fail. This is the case, for example, for Picard modular surfaces [associated with the group $U(2, 1)$] modulo a prime p which is inert in the underlying quadratic imaginary field. In this case, the abelian varieties parametrized by the open dense μ-ordinary stratum [30] are not ordinary. More generally, this happens for Shimura varieties associated with $U(n, m)$ if $n \neq m$, and p is inert ([16], Lemma 8.10). Another complication present in these examples is the fact that modular forms on $U(n, m)$ admit Fourier–Jacobi (FJ) expansions at the cusps, which are q-expansions with *theta functions* as coefficients.

One of the main goals of this paper is to define the theta operator for Picard modular surfaces at a good inert prime and study its properties. To explain how we overcome the need to consider the unit-root splitting of the cohomology of the universal abelian variety, let us re-examine the case of the modular curve X of full level $N \geq 3$ over $\mathbb{Z}_p$ ($(p, N) = 1$). We follow an approach of Gross [13], see also [1], who extended it to Hilbert modular varieties. Let κ be a fixed algebraic closure of $\mathbb{F}_p$, and consider the geometric characteristic p fiber X_κ. Let $\mathcal{A}$ be the universal elliptic curve over $Y = X\backslash C$ (the complement of the cusps) and let $\mathcal{L} = \omega_{\mathcal{A}/Y}$ be its cotangent bundle at the origin. Then $\mathcal{A}$ extends to a semi-abelian variety over X, and so does $\mathcal{L} = \omega_{\mathcal{A}/X}$. By definition, a weight k, level N modular form over κ, is a global section of $\mathcal{L}^k$ over X_κ, i.e.

$$M_k(N; \kappa) = H^0\left(X_\kappa, \mathcal{L}^k\right). \tag{0.2}$$

Let X_κ^{ord} be the ordinary locus in X_κ. Let $\tau : I \to X_\kappa^{\mathrm{ord}}$ be the Igusa curve of level p, classifying (besides the elliptic curve A and level structure classified by X_κ) embeddings of finite flat group schemes $\iota : \mu_p \hookrightarrow A[p]$. Let

$$h \in H^0\left(X_\kappa, \mathcal{L}^{p-1}\right) \tag{0.3}$$

be the Hasse invariant. As the universal $\iota : \mu_p \hookrightarrow \mathcal{A}[p]$ over I induces an isomorphism

$$\tau^*\mathcal{L} = \omega_{\mathcal{A}/I} = \omega_{\mathcal{A}[p]/I} \overset{\iota^*}{\simeq} \omega_{\mu_p/I} = \mathcal{O}_I, \tag{0.4}$$

the line bundle $\tau^*\mathcal{L}$ is trivialized over I by a canonical section a. In fact, $a^{p-1} = \tau^*h$.

Now, given a κ-valued modular form $f \in H^0(X_\kappa, \mathcal{L}^k)$, we consider its pull-back τ^*f to I, divide by a^k to get a function on I, and take its differential

$$\eta_f = \mathrm{d}\left(\tau^*f/a^k\right) \in \Omega_I^1. \tag{0.5}$$

The Gauss–Manin connection induces the Kodaira–Spencer isomorphism

$$\mathrm{KS} : \mathcal{L}^2 \otimes \mathcal{O}(C)^\vee \simeq \Omega_X^1. \tag{0.6}$$

As τ is étale, $\Omega_I^1 = \tau^* \Omega^1_{X_\kappa^{\mathrm{ord}}}$ and we may pull KS back to a similar isomorphism over I. We can therefore look at

$$a^k \cdot \mathrm{KS}^{-1}(\eta_f). \tag{0.7}$$

This is a section of $\tau^*(\mathcal{L}^{k+2} \otimes \mathcal{O}(C)^\vee)$ over I. Since we divided and multiplied by the same power of a, it descends to X_κ^{ord}. A calculation shows that it has at most *simple* poles at the supersingular points X_κ^{ss}, so

$$\theta(f) = h \cdot a^k \cdot \mathrm{KS}^{-1}(\eta_f) \tag{0.8}$$

extends to a global section of $\mathcal{L}^{k+p+1} \otimes \mathcal{O}(C)^\vee$ over X_κ, i.e. to a cusp form of weight $k+p+1$ and level N over κ. Note that $\theta(f)$ and $a^k \cdot \mathrm{KS}^{-1}(\eta_f)$ have the same q-expansions, since the q-expansion of h is 1. It can be checked that θ coincides with the operator denoted by $A\theta$ in [21].

The absence of the unit-root splitting from the above-mentioned construction can be "explained" by the use we made of the Igusa curve, which lies over the ordinary stratum. In the case of Picard modular surfaces at an inert prime p, it is nevertheless possible to construct an "Igusa surface" lying over the μ-ordinary part, even though the ordinary stratum (in the usual sense) is empty. Our construction of the theta operator is based on the same procedure, but there are now two automorphic vector bundles to consider, a line bundle $\mathcal{L}$ and a plane bundle $\mathcal{P}$. The Verschiebung homomorphism allows us to project the analogue of $\mathrm{KS}^{-1}(\eta_f)$ (which is a section of $\mathcal{P} \otimes \mathcal{L}$) to an appropriate one-dimensional piece.

The resulting operator Θ enjoys all the desired properties. It has the right effect on Fourier–Jacobi expansions, extends holomorphically across the 1-dimensional supersingular locus, and compares well with the theta operators on embedded modular curves. The theory of "theta cycles" [19] even presents a surprise (see 4.1).

Let us now review the contents of the paper in more detail. We denote by $\mathcal{K}$ a quadratic imaginary field and by $\bar{S}$ a compactified integral model of the Picard modular surface of full level $N \geq 3$, associated with $\mathcal{K}$. The surface $\bar{S}$ is defined over $R_0 = \mathcal{O}_{\mathcal{K}}[1/2D_{\mathcal{K}}N]$ and we may consider its reduction modulo the prime p, which is assumed to be relatively prime to $2N$ and inert in $\mathcal{K}$. For simplicity, fix an algebraic closure κ of R_0/pR_0 and consider the geometric fiber $\bar{S}_\kappa = \bar{S} \times_{\mathrm{Spec}(R_0)} \mathrm{Spec}(\kappa)$. Let $\mathcal{A}$ be the universal semi-abelian variety over $\bar{S}$. It is relatively 3-dimensional, has complex multiplication by $\mathcal{O}_{\mathcal{K}}$, and the cotangent bundle at the origin, $\omega_{\mathcal{A}/\bar{S}}$, is of type (2, 1). This means that if $\Sigma : \mathcal{O}_{\mathcal{K}} \hookrightarrow R_0$ is the canonical embedding and $\bar{\Sigma}$ its complex conjugate, then

$$\omega_{\mathcal{A}/\bar{S}} = \mathcal{P} \oplus \mathcal{L}, \tag{0.9}$$

where $\mathcal{P} = \omega_{\mathcal{A}/\bar{S}}(\Sigma)$ is a plane bundle on which $\mathcal{O}_{\mathcal{K}}$ acts via Σ, and $\mathcal{L} = \omega_{\mathcal{A}/\bar{S}}(\bar{\Sigma})$ is a line bundle on which it acts via $\bar{\Sigma}$. Scalar modular forms of weight $k \geq 0$ defined over an R_0-algebra R are by definition elements of

$$M_k(N; R) := H^0\left(\bar{S}_R, \mathcal{L}^k\right). \tag{0.10}$$

Our main interest is in $R = \kappa$. In this case, there are homomorphisms of vector bundles $V_{\mathcal{P}} : \mathcal{P} \to \mathcal{L}^{(p)}$ and $V_{\mathcal{L}} : \mathcal{L} \to \mathcal{P}^{(p)}$ deduced from the Verschiebung homomorphism. Here, for any vector bundle $\mathcal{V}$ over $\bar{S}_\kappa$, $\mathcal{V}^{(p)}$ stands for its base change under the absolute Frobenius morphism of degree p, $\Phi : \bar{S}_\kappa \to \bar{S}_\kappa$. The *Hasse invariant* is the map

$$h_{\bar{\Sigma}} = V_{\mathcal{P}}^{(p)} \circ V_{\mathcal{L}} : \mathcal{L} \to \mathcal{L}^{(p^2)}. \tag{0.11}$$

Since $\mathcal{L}$ is a line bundle, $\mathcal{L}^{(p)} \simeq \mathcal{L}^p$, so $h_{\bar{\Sigma}} \in H^0(\bar{S}_\kappa, \mathcal{L}^{p^2-1})$ is a modular form of weight p^2-1 over κ. The divisor of $h_{\bar{\Sigma}}$ is precisely the supersingular locus $S_{\rm ss} \subset \bar{S}_\kappa$. This is a reduced 1-dimensional closed subscheme whose geometric points x are characterized by the fact that $\mathcal{A}_x$ is supersingular (the Newton polygon of its p-divisible group has constant slope 1/2). The structure of $S_{\rm ss}$ has been determined by Vollaard [39], following work of Bültel and Wedhorn [4]. Its irreducible components are curves whose normalizations are all isomorphic to the Fermat curve of degree $p+1$. (If N is large enough, depending on p, these components are even non-singular.) They intersect transversally at finitely many points, which form the singular locus of $S_{\rm ss}$. This singular locus is also the superspecial locus $S_{\rm ssp}$ in $\bar{S}_\kappa$, characterized by the fact that $x \in S_{\rm ssp}$ if and only if $\mathcal{A}_x$ is isomorphic to a product of three supersingular elliptic curves. At $x \in S_{\rm ssp}$, the maps $V_{\mathcal{P}}$ and $V_{\mathcal{L}}$ vanish, but over the *general* supersingular locus $S_{\rm gss} = S_{\rm ss} \backslash S_{\rm ssp}$, they are both of rank 1. The complement of $S_{\rm ss}$ in $\bar{S}_\kappa$ is the dense, open μ-ordinary locus $\bar{S}_\mu$. Over a μ-ordinary point which does not belong to a cuspidal component, the p-divisible group of $\mathcal{A}_x$ is a product of a height 2 group of multiplicative type, a height 2 group of local–local type, and a height 2 étale group (all stable under $\mathcal{O}_\mathcal{K}$). See [8] and Sect. 2.1.2.

Section 1 is a rather thorough introduction to Picard modular surfaces and modular forms that will serve us also in future work. Occasionally (e.g. when we compute the Gauss–Manin connection in the complex model), we could not find a reference for the results in the form that was needed. We preferred to work them out from scratch, rather than embark on a tedious translation of notation. This section benefitted in several places from the excellent exposition in Bellaïche's thesis [2].

In Sect. 2 we review the geometry of $\bar{S}$ and the automorphic vector bundles $\mathcal{P}$ and $\mathcal{L}$ modulo an inert prime p. Here we follow [4,39], and the exposition in [8]. We construct the Igusa surface of level p. It is a finite étale Galois cover

$$\tau : \bar{I}g_\mu \to \bar{S}_\mu \tag{0.12}$$

of the μ-ordinary part in $\bar{S}_\kappa$, with Galois group $\Delta(p) = (\mathcal{O}_\mathcal{K}/p\mathcal{O}_\mathcal{K})^\times$. We prove that it is relatively irreducible, and compactify it over the supersingular locus to get a normal surface $\bar{I}g$, finite and flat over $\bar{S}_\kappa$, which is totally ramified over $S_{\rm ss}$. The Hasse invariant has a tautological p^2-1 root a over the whole of $\bar{I}g$. Thus $a \in H^0(\bar{I}g, \tau^*\mathcal{L})$ and $a^{p^2-1} = \tau^* h_{\bar{\Sigma}}$.

In Sect. 3 we construct the theta operator. We pull back $f \in H^0(\bar{S}_\kappa, \mathcal{L}^k)$ to $\bar{I}g_\mu$, divide by the non-vanishing section a^k to get a function, and let

$$\eta_f = \mathrm{d}\left(\tau^* f/a^k\right) \in H^0\left(\bar{I}g_\mu, \Omega^1\right). \tag{0.13}$$

The Kodaira–Spencer isomorphism over S is an isomorphism of rank two vector bundles

$$\mathrm{KS} : \mathcal{P} \otimes \mathcal{L} \simeq \Omega^1_S. \tag{0.14}$$

When we try to extend it to $\bar{S}$, we find out that it has a pole along the cuspidal divisor $C = \bar{S} \backslash S$. Nevertheless, in the characteristic p fiber, the map

$$(V_{\mathcal{P}} \otimes 1) \circ \mathrm{KS}^{-1} : \Omega^1_{S_\kappa} \to \mathcal{L}^{(p)} \otimes \mathcal{L} = \mathcal{L}^{p+1} \tag{0.15}$$

extends holomorphically across C, and even acquires a simple zero there. We pull it back from $\bar{S}_\mu$ to $\bar{I}g_\mu$ under the étale map τ, and define

$$\Theta(f) = a^k \cdot (V_{\mathcal{P}} \otimes 1) \circ \mathrm{KS}^{-1}(\eta_f) \in H^0\left(\bar{I}g_\mu, \tau^* \mathcal{L}^{k+p+1}\right). \tag{0.16}$$

Thanks to the fact that we have multiplied by a^k, this section descends to $\bar{S}_\mu$. A pleasant computation reveals that $\Theta(f)$ has no poles along S_{ss}. We end up with

$$\Theta(f) \in H^0\left(\bar{S}_\kappa, \mathcal{L}^{k+p+1}\right), \tag{0.17}$$

a weight $k+p+1$, level N modular form over κ.

It is curious to note that in the case of modular curves, $a^k \cdot \mathrm{KS}^{-1}(\eta_f)$ was of weight $k+2$, but had poles at the supersingular points, and only $\theta(f) = h \cdot a^k \cdot \mathrm{KS}^{-1}(\eta_f)$ extended holomorphically to a weight $k+p+1$ modular form. Here, the projection $V_{\mathcal{P}}$ takes care of the shift by $p+1$ in the weight, and at the same time reduces the order of the pole along $Ig_{\mathrm{ss}} = \bar{I}g \backslash \bar{I}g_\mu$, so that $\Theta(f)$ becomes holomorphic over the whole surface.

The ultimate justification for our construction comes when we compute the effect of Θ on Fourier–Jacobi expansions, which is essentially a "Tate twist". The computation uses both p-adic and complex formalisms. It may be possible to perform it entirely on the "Mumford-Tate object" (see Section 4.5 of [9,26]), but we believe that our approach has its own didactical merit.

In Sect. 4 we compare our theta operator with theta operators on embedded modular curves. We also discuss theta cycles and filtrations on modular forms mod p.

Section 5 brings up p-adic modular forms in the style of Serre and Katz. The study of overconvergent forms, intimately connected with the study of the canonical subgroup and Coleman's classicality theorem, will be the subject of another paper.

Many of the results of this paper, including the construction of the theta operator, generalize to unitary Shimura varieties associated with $U(n-1,1)$ for general n. Another direction in which the setup could be generalized is to replace $\mathcal{K}$ by an arbitrary CM field. This seems to require substantial additional work, apart from a heavy load of notation, even if the general lay-out would be the same. We refer the reader to [18] for a detailed discussion of some of the topics treated here over general CM fields (albeit for a split prime p).

1 Background

1.1 The unitary group and its symmetric space

1.1.1 Notation

Let $\mathcal{K}$ be an imaginary quadratic field, contained in $\mathbb{C}$. We denote by $\Sigma : \mathcal{K} \hookrightarrow \mathbb{C}$ the inclusion and by $\bar{\Sigma} : \mathcal{K} \hookrightarrow \mathbb{C}$ its complex conjugate. We use the following notation:

- $d_{\mathcal{K}}$—the square-free integer such that $\mathcal{K} = \mathbb{Q}(\sqrt{d_{\mathcal{K}}})$.
- $D_{\mathcal{K}}$—the discriminant of $\mathcal{K}$, equal to $d_{\mathcal{K}}$ if $d_{\mathcal{K}} \equiv 1 \mod 4$ and $4d_{\mathcal{K}}$ if $d_{\mathcal{K}} \equiv 2, 3 \mod 4$.
- $\delta_{\mathcal{K}} = \sqrt{D_{\mathcal{K}}}$—the square root with positive imaginary part, a generator of the different of $\mathcal{K}$, sometimes simply denoted δ.
- $\omega_{\mathcal{K}} = (1+\sqrt{d_{\mathcal{K}}})/2$ if $d_{\mathcal{K}} \equiv 1 \mod 4$, otherwise $\omega_{\mathcal{K}} = \sqrt{d_{\mathcal{K}}}$, so that $\mathcal{O}_{\mathcal{K}} = \mathbb{Z} + \mathbb{Z}\omega_{\mathcal{K}}$.
- $\bar{a}$—the complex conjugate of $a \in \mathcal{K}$.
- $\mathrm{Im}_\delta(a) = (a - \bar{a})/\delta$, for $a \in \mathcal{K}$.

We fix an integer $N \geq 3$ (the "tame level") and let $R_0 = \mathcal{O}_{\mathcal{K}}[1/(2d_{\mathcal{K}}N)]$. This is our *base ring*. If R is any R_0-algebra and M is any R-module with $\mathcal{O}_{\mathcal{K}}$-action, then M becomes an $\mathcal{O}_{\mathcal{K}} \otimes R$-module and we have a canonical *type decomposition*

$$M = M(\Sigma) \oplus M(\bar{\Sigma}) \tag{1.1}$$

where $M(\Sigma) = e_\Sigma M$ and $M(\bar{\Sigma}) = e_{\bar{\Sigma}} M$, and where the idempotents e_Σ and $e_{\bar{\Sigma}}$ are defined by

$$e_\Sigma = \frac{1\otimes 1}{2} + \frac{\delta\otimes\delta^{-1}}{2}, \quad e_{\bar{\Sigma}} = \frac{1\otimes 1}{2} - \frac{\delta\otimes\delta^{-1}}{2}. \tag{1.2}$$

Then $M(\Sigma)$ (resp. $M(\bar{\Sigma})$) is the part of M on which $\mathcal{O}_\mathcal{K}$ acts via Σ (resp. $\bar{\Sigma}$). The same notation will be used for sheaves of modules on R-schemes, endowed with an $\mathcal{O}_\mathcal{K}$ action. If M is locally free, we say that it has *type* (p, q) if $M(\Sigma)$ is of rank p and $M(\bar{\Sigma})$ is of rank q.

We denote by

$$\mathbf{T} = \mathrm{res}_{\mathbb{Q}}^{\mathcal{K}} \mathbb{G}_m \tag{1.3}$$

the non-split torus whose $\mathbb{Q}$-points are $\mathcal{K}^\times$, and by ρ the non-trivial automorphism of $\mathbf{T}$, which on $\mathbb{Q}$-points induces $\rho(a) = \bar{a}$. The group $\mathbb{G}_m$ embeds in $\mathbf{T}$ and the homomorphism $a \mapsto a\cdot\rho(a)$ from $\mathbf{T}$ to itself factors through a homomorphism

$$N : \mathbf{T} \to \mathbb{G}_m, \tag{1.4}$$

the *norm homomorphism*. Its kernel $\ker(N)$ is denoted $\mathbf{T}^1$.

1.1.2 The unitary group

Let $V = \mathcal{K}^3$ and endow it with the hermitian pairing

$$(u, v) = {}^t\bar{u} \begin{pmatrix} & & \delta^{-1} \\ & 1 & \\ -\delta^{-1} & & \end{pmatrix} v. \tag{1.5}$$

We identify $V_\mathbb{R}$ with $\mathbb{C}^3$ ($\mathcal{K}$ acting via the natural inclusion Σ). It then becomes a hermitian space of signature (2, 1). Conversely, any 3-dimensional hermitian space over $\mathcal{K}$ whose signature at the infinite place is (2, 1) is isomorphic to V after rescaling the hermitian form by a positive rational number.

Let

$$\mathbf{G} = \mathbf{GU}(V, (,)) \tag{1.6}$$

be the general unitary group of V, regarded as an algebraic group over $\mathbb{Q}$. For any $\mathbb{Q}$-algebra A, we have

$$\mathbf{G}(A) = \left\{(g, \mu) \in GL_3(A\otimes\mathcal{K}) \otimes A^\times |\ (gu, gv) = \mu\cdot(u, v)\ \forall u, v \in V_A\right\}. \tag{1.7}$$

We write $G = \mathbf{G}(\mathbb{Q})$, $G_\infty = \mathbf{G}(\mathbb{R})$ and $G_p = \mathbf{G}(\mathbb{Q}_p)$. A similar notational convention will apply to any algebraic group over $\mathbb{Q}$ without further ado. If p splits in $\mathcal{K}$, $\mathbb{Q}_p \otimes \mathcal{K} \simeq \mathbb{Q}_p^2$ and G_p becomes isomorphic to $GL_3(\mathbb{Q}_p) \times \mathbb{Q}_p^\times$. The isomorphism depends on the embedding of $\mathcal{K}$ in $\mathbb{Q}_p$, i.e. on the choice of a prime above p in $\mathcal{K}$. For a non-split prime p, the group G_p, like G_∞, is of (semisimple) rank 1.

As μ is determined by g, we often abuse notation and write g for the pair (g, μ) and $\mu(g)$ for the *similitude factor (multiplier)* μ. It is a character of algebraic groups over $\mathbb{Q}$, $\mu : \mathbf{G} \to \mathbb{G}_m$. Another character is $\det : \mathbf{G} \to \mathbf{T}$, defined by $\det(g, \mu) = \det(g)$. If we let

$$\nu = \mu^{-1} \cdot \det : \mathbf{G} \to \mathbf{T} \tag{1.8}$$

then both μ and det are expressible in terms of ν, namely $\mu = \nu\cdot(\rho\circ\nu)$ and $\det = \nu^2\cdot(\rho\circ\nu)$.

The groups

$$\mathbf{U} = \ker\mu, \quad \mathbf{SU} = \ker\nu = \ker\mu \cap \ker(\det) \tag{1.9}$$

are the unitary and the special unitary group, respectively.

We also introduce an alternating $\mathbb{Q}$-linear pairing $\langle , \rangle : V \times V \to \mathbb{Q}$ (the *polarization form*) defined by $\langle u, v \rangle = \mathrm{Im}_\delta(u, v)$. We then have the formulae

$$\langle au, v \rangle = \langle u, \bar{a}v \rangle , \quad 2(u, v) = \langle u, \delta v \rangle + \delta \langle u, v \rangle . \tag{1.10}$$

1.1.3 The hermitian symmetric domain

The group $G_\infty = \mathbf{G}(\mathbb{R})$ acts on $\mathbb{P}^2_{\mathbb{C}} = \mathbb{P}(V_{\mathbb{R}})$ by projective linear transformations and preserves the open subdomain $\mathfrak{X}$ of negative definite lines (in the metric $(,)$), which is biholomorphic to the open unit ball in $\mathbb{C}^2$. Every negative definite line is represented by a unique vector ${}^t(z, u, 1)$, and such a vector represents a negative definite line if and only if

$$\lambda(z, u) \overset{\mathrm{def}}{=} \mathrm{Im}_\delta(z) - u\bar{u} > 0. \tag{1.11}$$

One refers to the realization of $\mathfrak{X}$ as the set of points $(z, u) \in \mathbb{C}^2$ satisfying this inequality as a *Siegel domain of the second kind*. It is convenient to think of the point $x_0 = (\delta/2, 0)$ as the "center" of $\mathfrak{X}$.

If we let K_∞ be the stabilizer of x_0 in G_∞, then K_∞ is compact modulo center ($K_\infty \cap \mathbf{U}(\mathbb{R})$ is compact and isomorphic to $U(2) \times U(1)$). Since G_∞ acts transitively on $\mathfrak{X}$, we may identify $\mathfrak{X}$ with G_∞/K_∞.

The usual upper half plane embeds in $\mathfrak{X}$ as the set of points where $u = 0$.

1.1.4 The cusps of $\mathfrak{X}$

The boundary $\partial\mathfrak{X}$ of $\mathfrak{X}$ is the set of points (z, u) where $\mathrm{Im}_\delta(z) = u\bar{u}$, together with a unique point "at infinity" c_∞ represented by the line ${}^t(1 : 0 : 0)$. The lines represented by $\partial\mathfrak{X}$ are the isotropic lines in $V_{\mathbb{R}}$. The set of *cusps* $\mathcal{C}\mathfrak{X}$ is the set of $\mathcal{K}$-rational isotropic lines. If $s \in \mathcal{K}$ and $r \in \mathbb{Q}$, we write

$$c_s^r = (r + \delta s\bar{s}/2, s). \tag{1.12}$$

Then $\mathcal{C}\mathfrak{X} = \left\{c_s^r | r \in \mathbb{Q}, s \in \mathcal{K}\right\} \cup \{c_\infty\}$. The group $G = \mathbf{G}(\mathbb{Q})$ acts transitively on the cusps.

The stabilizer of a cusp is a Borel subgroup in G_∞. Since G acts transitively on the cusps, we may assume that our cusp is c_∞. It is then easy to check that its stabilizer P_∞ has the form $P_\infty = M_\infty N_\infty$, where

$$M_\infty = \left\{ tm(\alpha, \beta) = t \begin{pmatrix} \alpha & & \\ & \beta & \\ & & \bar{\alpha}^{-1} \end{pmatrix} | \; t \in \mathbb{R}_+^\times, \alpha \in \mathbb{C}^\times, \beta \in \mathbb{C}^1 \right\}, \tag{1.13}$$

$$N_\infty = \left\{ n(u, r) = \begin{pmatrix} 1 & \delta\bar{u} & r + \delta u\bar{u}/2 \\ & 1 & u \\ & & 1 \end{pmatrix} | \; u \in \mathbb{C}, r \in \mathbb{R} \right\}. \tag{1.14}$$

The matrix $tm(\alpha, \beta)$ belongs to U_∞ if and only if $t = 1$, and to SU_∞ if furthermore $\beta = \bar{\alpha}/\alpha$. The group N_∞ is contained in SU_∞. Since $N = N_\infty \cap G$ still acts transitively on the set of finite cusps c_s^r, we conclude that G acts *doubly transitively* on $\mathcal{C}\mathfrak{X}$.

Of particular interest to us will be the *geodesics* connecting an interior point (z, u) to a cusp $c \in \mathcal{C}\mathfrak{X}$. If $(z, u) = n(u, r)m(d, 1)x_0$ (recall $x_0 = {}^t(\delta/2 : 0 : 1)$) where d is real and positive (i.e. $r = \Re z$ and $d = \sqrt{\lambda(z, u)}$), then the geodesic connecting (z, u) to c_∞ can be described by the formula

$$\begin{aligned} \gamma_u^r(t) &= n(u, r)m(t, 1)x_0 \\ &= (r + \delta(u\bar{u} + t^2)/2, u) \;\; (d \le t < \infty). \end{aligned} \tag{1.15}$$

The same geodesic extends in the opposite direction for $0 < t \leq d$, and if u and r lie in $\mathcal{K}$, it ends there in the cusp c_u^r. We shall call $\gamma_u^r(t)$ the *geodesic retraction* of $\mathfrak{X}$ to the cusp c_∞. As $0 < t < \infty$, these parallel geodesics exhaust $\mathfrak{X}$.

1.2 Picard modular surfaces over $\mathbb{C}$

1.2.1 Lattices and their arithmetic groups

Fix an $\mathcal{O}_\mathcal{K}$-invariant lattice $L \subset V$ which is *self-dual* in the sense that

$$L = \{u \in V |\ \langle u, v\rangle \in \mathbb{Z}\ \forall v \in L\}. \tag{1.16}$$

Equivalently, L is its own $\mathcal{O}_\mathcal{K}$-dual with respect to the hermitian pairing $(,)$. We assume also that the Steinitz class[1] of L as an $\mathcal{O}_\mathcal{K}$-module is $[\mathcal{O}_\mathcal{K}]$, or, what amounts to the same, that L is a free $\mathcal{O}_\mathcal{K}$-module. When we introduce the Shimura variety later on, we shall relax this last assumption, but the resulting scheme will be disconnected (over $\mathbb{C}$).

Fix an integer $N \geq 1$ and let

$$\Gamma = \Gamma_L(N) = \left\{g \in G |\ gL = L \text{ and } g(u) \equiv u \mod NL\ \forall u \in L\right\}, \tag{1.17}$$

a discrete subgroup of G_∞. It is easy to see that if $N \geq 3$, then Γ is torsion free, acts freely and faithfully on $\mathfrak{X}$, and is contained in SU_∞. From now on, we assume that this is the case.

If $g \in G$ and $\mu(g) = 1$ (i.e. $g \in U$), the lattice gL is another lattice of the same sort and the discrete group corresponding to it is $g\Gamma g^{-1}$. Since U acts transitively on the cusps, this reduces the study of $\Gamma\backslash\mathfrak{X}$ near a cusp to the study of a neighborhood of the standard cusp c_∞ (at the price of changing L and Γ).

It is important to know the classification of lattices L as above (self-dual and $\mathcal{O}_\mathcal{K}$-free). Let e_1, e_2, e_3 be the standard basis of $\mathcal{K}^3$. Let

$$L_0 = \mathrm{Span}_{\mathcal{O}_\mathcal{K}}\{\delta e_1, e_2, e_3\} \tag{1.18}$$

and

$$L_1 = \mathrm{Span}_{\mathcal{O}_\mathcal{K}}\left\{\frac{\delta}{2}e_1 + e_3, e_2, \frac{\delta}{2}e_1 - e_3\right\}. \tag{1.19}$$

These two lattices are self-dual and, of course, $\mathcal{O}_\mathcal{K}$-free. The following theorem is based on the local–global principle and a classification of lattices over $\mathbb{Q}_p$ by Shimura [35].

Lemma 1.1 ([28], p. 25) *For any lattice L as above there exists a $g \in U$ such that $gL = L_0$ or $gL = L_1$. If $D_\mathcal{K}$ is odd, L_0 and L_1 are equivalent. If $D_\mathcal{K}$ is even, they are inequivalent.*

Indeed, if $D_\mathcal{K}$ is even, $L_0 \otimes \mathbb{Q}_p$ and $L_1 \otimes \mathbb{Q}_p$ are U_p-equivalent for every $p \neq 2$, but not for $p = 2$.

1.2.2 Picard modular surfaces and the Baily–Borel compactification

We denote by X_Γ the complex surface $\Gamma\backslash\mathfrak{X}$. Since the action of Γ is free, X_Γ is smooth. We describe a topological compactification of X_Γ. A *standard neighborhood* of the cusp c_∞ in $\mathfrak{X}$ is an open set of the form

$$\Omega_R = \left\{(z, u) | \lambda(z, u) > R\right\}. \tag{1.20}$$

The set $\mathcal{C}_\Gamma = \Gamma\backslash\mathcal{C}\mathfrak{X}$ is finite, and we write $c_\Gamma = \Gamma c$. We let X_Γ^* be the disjoint union of X_Γ and $\mathcal{C}_\Gamma$. Let Γ_{cusp} be the stabilizer of c_∞ in Γ. We topologize X_Γ^* by taking

[1] The Steinitz class of a finite projective $\mathcal{O}_\mathcal{K}$-module is the class of its top exterior power as an invertible module.

$\Gamma_{\mathrm{cusp}}\backslash\Omega_R \cup \{c_{\infty,\Gamma}\}$ as a basis of neighborhoods at $c_{\infty,\Gamma}$. If $c = g(c_\infty)$ where $g \in U$, we take $g(g^{-1}\Gamma_{\mathrm{cusp}}g\backslash\Omega_R) \cup \{c_\Gamma\}$ instead. The following theorem is well known.

Theorem 1.2 (Satake, Baily–Borel) *X_Γ^* is projective, and the singularities at the cusps are normal. In other words, there exists a normal complex projective surface S_Γ^* and a homeomorphism $\iota : S_\Gamma^*(\mathbb{C}) \simeq X_\Gamma^*$, which on $S_\Gamma(\mathbb{C}) = \iota^{-1}(X_\Gamma)$ is an isomorphism of complex manifolds. S_Γ^* is uniquely determined up to isomorphism.*

1.2.3 The universal abelian variety over X_Γ

With $x \in \mathfrak{X}$ and with our choice of L, we shall now associate a PEL structure $\underline{A}_x = (A_x, \lambda_x, \iota_x, \alpha_x)$ where

(1) A_x is a 3-dimensional complex abelian variety,
(2) λ_x is a principal polarization on A_x (i.e. an isomorphism $A_x \simeq A_x^t$ with its dual abelian variety induced by an ample line bundle),
(3) $\iota_x : \mathcal{O}_\mathcal{K} \hookrightarrow End(A_x)$ is an embedding of CM type (2, 1) (i.e. the action of $\iota(a)$ on the tangent space of A_x at the origin induces the representation $2\Sigma + \bar{\Sigma}$) such that the Rosati involution induced by λ_x preserves $\iota(\mathcal{O}_\mathcal{K})$ and is given by $\iota(a) \mapsto \iota(\bar{a})$,
(4) $\alpha_x : N^{-1}L/L \simeq A_x[N]$ is a full level N structure, compatible with the $\mathcal{O}_\mathcal{K}$-action and the polarization. The latter condition means that if we denote by $\langle , \rangle_\lambda$ the Weil "e_N-pairing" on $A_x[N]$ induced by λ_x, then for $l, l' \in N^{-1}L$

$$\left\langle \alpha_x(l), \alpha_x(l')\right\rangle_\lambda = e^{2\pi i N\langle l,l'\rangle}. \tag{1.21}$$

Let W_x be the negative definite complex line in $V_\mathbb{R} = \mathbb{C}^3$ defined by x, and $W_x^\perp$ its orthogonal complement, a positive definite plane. Let J_x be the complex structure which is multiplication by i on $W_x^\perp$ and by $-i$ on W_x. Let $A_x = (V_\mathbb{R}, J_x)/L$. Then the polarization form $\langle , \rangle$ is a Riemann form on L, which determines a principal polarization on A_x as usual. The action of $\mathcal{O}_\mathcal{K}$ is derived from the underlying $\mathcal{K}$ structure of V. As we have reversed the complex structure on W_x, the CM type is now (2,1). Finally the level N structure α_x is the identity map.

If $\gamma \in \Gamma$, then γ induces an isomorphism between $\underline{A}_x$ and $\underline{A}_{\gamma(x)}$. Conversely, if $\underline{A}_x$ and $\underline{A}_{x'}$ are isomorphic structures, it is easy to see that x' and x must belong to the same Γ-orbit. It follows that points of X_Γ are in a bijection with PEL structures of the above type for which the triple

$$(H_1(A_x, \mathbb{Z}), \iota_x, \langle , \rangle_{\lambda_x}) \tag{1.22}$$

is isomorphic to $(L, \iota, \langle , \rangle)$ (here ι refers to the $\mathcal{O}_\mathcal{K}$ action on L), with the further condition that α_x is compatible with the isomorphism between L and $H_1(A_x, \mathbb{Z})$ in the sense that we have a commutative diagram

$$\begin{array}{ccccccc} 0 \to & L & \to & N^{-1}L & \to & N^{-1}L/L & \to 0 \\ & \downarrow & & \downarrow & & \downarrow \alpha_x & \\ 0 \to & H_1(A_x, \mathbb{Z}) & \to & N^{-1}H_1(A_x, \mathbb{Z}) & \to & A_x[N] & \to 0 \end{array} . \tag{1.23}$$

1.2.4 A "moving lattice" model for the universal abelian variety

We want to assemble the individual A_x into an *abelian variety A over $\mathfrak{X}$*. In other words, we want to construct a 5-dimensional complex manifold A, together with a holomorphic map $A \to \mathfrak{X}$ whose fiber over x is identified with A_x. For that, as well as for the computation of the Gauss–Manin connection below, it is convenient to introduce another model, in which the complex structure on $\mathbb{C}^3$ is fixed, but the lattice varies.

For simplicity, we assume from now on that $L = L_0$ is spanned over $\mathcal{O}_\mathcal{K}$ by $\delta e_1, e_2$ and e_3. The case of L_1 can be handled similarly.

Let $\mathbb{C}^3$ be given the usual complex structure, and let $a \in \mathcal{O}_\mathcal{K}$ act on it via the matrix

$$\iota'(a) = \begin{pmatrix} a & & \\ & a & \\ & & \bar{a} \end{pmatrix}. \tag{1.24}$$

Given $x = (z, u) \in \mathfrak{X}$, consider the lattice

$$L'_x = \mathrm{Span}_{\iota'(\mathcal{O}_\mathcal{K})} \left\{ \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ -u \end{pmatrix}, \begin{pmatrix} u \\ -z/\delta \\ z/\delta \end{pmatrix} \right\} \subset \mathbb{C}^3. \tag{1.25}$$

The map $T_x : \mathbb{C}^3 \to \mathbb{C}^3$ which sends $\zeta = {}^t(\zeta_1, \zeta_2, \zeta_3)$ to

$$T(\zeta) = \lambda(z, u)^{-1} \left\{ -\zeta_1 \begin{pmatrix} \bar{u}z \\ (z - \bar{z})/\delta \\ \bar{u} \end{pmatrix} - \zeta_2 \begin{pmatrix} \bar{z} + \delta u\bar{u} \\ u \\ 1 \end{pmatrix} + \bar{\zeta}_3 \begin{pmatrix} z \\ u \\ 1 \end{pmatrix} \right\} \tag{1.26}$$

is a complex linear isomorphism between $\mathbb{C}^3$ and $(V_\mathbb{R}, J_x)$. In fact, it sends $\mathbb{C}e_1 + \mathbb{C}e_2$ linearly to $W_x^\perp$ and $\mathbb{C}e_3$ conjugate-linearly to W_x. It intertwines the ι' action of $\mathcal{O}_\mathcal{K}$ on $\mathbb{C}^3$ with its ι-action on $(V_\mathbb{R}, J_x)$. It furthermore sends L'_x to L. In fact, an easy computation shows that it sends the three generating vectors of L'_x to $\delta e_1, e_2$ and e_3, respectively. We conclude that T_x induces an isomorphism

$$T_x : A'_x = \mathbb{C}^3/L'_x \simeq A_x. \tag{1.27}$$

Consider the differential forms $\mathrm{d}\zeta_1, \mathrm{d}\zeta_2$ and $\mathrm{d}\zeta_3$. As their periods along any $l \in L'_x$ vary holomorphically in z and u, the five coordinates $\zeta_1, \zeta_2, \zeta_3, z, u$ form a local system of coordinates on the family $A' \to \mathfrak{X}$. Identifying A' with A allows us to put the desired complex structure on the family A. Alternatively, we may define A' as the quotient of $\mathbb{C}^3 \times \mathfrak{X}$ by $\zeta \mapsto \zeta + l(z, u)$ where $l(z, u)$ varies over the holomorphic lattice-sections.

The model A' has another advantage that will become clear when we examine the degeneration of the universal abelian variety at the cusp c_∞. It suffices to note at this point that the first two of the three generating vectors of L'_x depend only on u.

1.3 The Picard moduli scheme

1.3.1 The moduli problem

Fix $N \geq 3$ and $L = L_0 \subset V$ as before. Let R be an R_0-algebra. Let $\mathcal{M}(R)$ be the collection of (isomorphism classes of) PEL structures $(A, \lambda, \iota, \alpha)$ where

(1) A/R is an abelian scheme of relative dimension 3
(2) $\lambda : A \simeq A^t$ is a principal polarization
(3) $\iota : \mathcal{O}_\mathcal{K} \to End(A/R)$ is a homomorphism such that (1) ι makes $Lie(A/R)$ a locally free R-module of type (2, 1), (2) the Rosati involution induced on $\iota(\mathcal{O}_\mathcal{K})$ by λ is $\iota(a) \mapsto \iota(\bar{a})$.
(4) $\alpha : N^{-1}L/L \simeq A[N]$ is an isomorphism of $\mathcal{O}_\mathcal{K}$-group schemes over R which is compatible with the polarization in the sense that there exists an isomorphism $\nu_N : \mathbb{Z}/N\mathbb{Z} \simeq \mu_N$ of group schemes over R such that

$$\left\langle \alpha\left(\frac{l}{N}\right), \alpha\left(\frac{l'}{N}\right) \right\rangle_\lambda = \nu_N\left(\langle l, l' \rangle \mod N\right). \tag{1.28}$$

In addition we require that for every multiple N' of N, locally étale over $\mathrm{Spec}(R)$, there exists a similar level N'-structure α', restricting to α on $N^{-1}L/L$. One says that α is *locally étale symplectic liftable* ([26], 1.3.6.2).

In view of Lemma 1.1, the last condition of symplectic liftability is void if $D_{\mathcal{K}}$ is odd, while if $D_{\mathcal{K}}$ is even it is equivalent to the following condition ([2], I.3.1):

- For any geometric point $\eta : R \to k$ (k algebraically closed field, necessarily of characteristic different from 2), the $\mathcal{O}_{\mathcal{K}} \otimes \mathbb{Z}_2$ polarized module $(T_2 A_\eta, \langle,\rangle_\lambda)$ is isomorphic to $(L \otimes \mathbb{Z}_2, \langle,\rangle)$ under a suitable identification of $\lim_{\leftarrow} \mu_{2^n}(k)$ with $\mathbb{Z}_2$.

The choice of L_0 was arbitrary. If we took L_1 as our basic lattice, we would get a similar moduli problem.

A level N structure α can exist only if the group schemes $\mathbb{Z}/N\mathbb{Z}$ and μ_N become isomorphic over R, but the isomorphism ν_N is then determined by α.

$\mathcal{M}$ becomes a functor on the category of R_0-algebras (and more generally, on the category of R_0-schemes) in the obvious way. The following theorem is of fundamental importance ([26], I.4.1.11).

Theorem 1.3 *The functor $R \mapsto \mathcal{M}(R)$ is represented by a smooth quasi-projective scheme S over* $\mathrm{Spec}(R_0)$, *of relative dimension 2.*

We call S the *(open) Picard modular surface of level N*. It comes equipped with a universal structure $(\mathcal{A}, \lambda, \iota, \alpha)$ of the above type over S. We call $\mathcal{A}$ the *universal abelian scheme* over S. For every R_0-algebra R and PEL structure in $\mathcal{M}(R)$, there exists a unique R-point of S such that the given PEL structure is obtained from the universal one by base change.

We refer to [26], 1.4.3 for the relation between the given formulation of the moduli problem and other formulations due, e.g. to Kottwitz.

1.3.2 The Shimura variety Sh_K

We briefly recall the interpretation of the Picard modular surface as a canonical model of a Shimura variety. The symmetric domain $\mathfrak{X}$ can be interpreted as a G_∞-conjugacy class of homomorphisms

$$h : \mathbb{S} = \mathrm{res}_{\mathbb{R}}^{\mathbb{C}} \mathbb{G}_m \to \mathbf{G}_{\mathbb{R}} \tag{1.29}$$

turning $(\mathbf{G}, \mathfrak{X})$ into a Shimura datum in the sense of Deligne [7]. In fact $h_x(i) = J_x$. The reflex field associated with this datum turns out to be $\mathcal{K}$. Let K_∞ be the stabilizer of x_0 in G_∞ and $K_f^0 \subset \mathbf{G}(\mathbb{A}_f)$ the subgroup stabilizing $\widehat{L} = L \otimes \widehat{\mathbb{Z}}$. Let K_f be the subgroup of K_f^0 inducing the identity on L/NL. Let $K = K_\infty K_f \subset \mathbf{G}(\mathbb{A})$. Then the Shimura variety Sh_K is a complex quasi-projective variety whose complex points are isomorphic, as a complex manifold, to the double coset space

$$\begin{aligned} \mathrm{Sh}_K(\mathbb{C}) &\simeq \mathbf{G}(\mathbb{Q})\backslash \mathbf{G}(\mathbb{A})/K \\ &= \mathbf{G}(\mathbb{Q})\backslash(\mathfrak{X} \times \mathbf{G}(\mathbb{A}_f)/K_f). \end{aligned} \tag{1.30}$$

The theory of Shimura varieties provides a *canonical model* for Sh_K over $\mathcal{K}$. The following important theorem complements the one on the representability of the functor $\mathcal{M}$.

Theorem 1.4 *The canonical model of Sh_K is the generic fiber $S_{\mathcal{K}}$ of S.*

Let us explain only how to associate with a point of $\mathrm{Sh}_K(\mathbb{C})$ a point in $S(\mathbb{C})$. For that we have to associate an element of $\mathcal{M}(\mathbb{C})$ with $g \in \mathbf{G}(\mathbb{A})$, and show that the structures associated with g and to γgk ($\gamma \in G, k \in K$) are isomorphic. Let $x = x_g = g_\infty(x_0) \in \mathfrak{X}$. Let $L_g = g_f(\widehat{L}) \cap V$ (the intersection taking place in $V_{\mathbb{A}} = \widehat{L} \otimes \mathbb{Q}$) and

$$A_g = (V_{\mathbb{R}}, J_x)/L_g. \tag{1.31}$$

Note that J_x depends only on $g_\infty K_\infty$ and L_g only on $g_f K_f^0$, so A_g depends only on gK^0.

Let $\tilde{\mu}(g)$ be the unique positive rational number such that for every prime p,

$$\mathrm{ord}_p \tilde{\mu}(g) = \mathrm{ord}_p \mu(g_p). \tag{1.32}$$

Such a rational number exists since $\mu(g_p)$ is a p-adic unit for almost all p and $\mathbb{Q}$ has class number 1. We claim that

$$\langle , \rangle_g = \tilde{\mu}(g)^{-1} \langle , \rangle : L_g \times L_g \to \mathbb{Q}$$

induces a principal polarization λ_g on A_g. That this is a (rational) Riemann form follows from the fact that $(u, v)_{J_x} = \langle u, J_x v\rangle + i \langle u, v\rangle$ is hermitian positive definite. That $\langle , \rangle_g$ is indeed $\mathbb{Z}$-valued and L_g is self-dual follows from the choice of $\tilde{\mu}(g)$ since locally at p the dual of $g_p L_p$ under $\langle , \rangle : V_p \times V_p \to \mathbb{Q}_p$ is $\mu(g_p)^{-1} g_p L_p$. We conclude that there exists a unique polarization $\lambda_g : A_g \to A_g^t$ such that

$$\langle u, v\rangle_{\lambda_g} = \exp(2\pi i l \langle u, v\rangle_g) \tag{1.33}$$

for every $u, v \in A_g[l] = l^{-1}L_g/L_g$ and every $l \geq 1$. This polarization is principal.

Since g_f commutes with the $\mathcal{K}$-structure on $V_{\mathbb{A}}$, L_g is still an $\mathcal{O}_{\mathcal{K}}$-lattice, hence ι_g is defined.

Finally α_g is derived from

$$N^{-1}L/L = N^{-1}\widehat{L}/\widehat{L} \xrightarrow{g_f} N^{-1}\widehat{L}_g/\widehat{L}_g = N^{-1}L_g/L_g = A_g[N]. \tag{1.34}$$

We note that α_g depends only on gK because $K_f \subset K_f^0$ is the principal level-N subgroup, and that it lifts to level N' structure for any multiple N' of N, by the same formula. The isomorphism $\nu_{N,g}$ between $\mathbb{Z}/N\mathbb{Z}$ and $\mu_N(\mathbb{C})$ that makes (1.28) work is self-evident [see (1.49)]. Let $\underline{A}_g \in \mathcal{M}(\mathbb{C})$ be the structure just constructed.

Let now $\gamma \in \mathbf{G}(\mathbb{Q})$. Then the action of γ on V induces an isomorphism between the tuples $\underline{A}_g$ and $\underline{A}_{\gamma g}$. Indeed, $\gamma : V_{\mathbb{R}} \to V_{\mathbb{R}}$ intertwines the complex structures x_g and $x_{\gamma g}$, and carries L_g to $L_{\gamma g}$, so induces an isomorphism of the abelian varieties, which clearly commutes with the PEL structures.

This shows that $\underline{A}_g$ depends solely on the double coset of g in $\mathbf{G}(\mathbb{Q})\backslash\mathbf{G}(\mathbb{A})/K$. One is left now with two tasks which we leave out: (i) proving that if $\underline{A}_g \simeq \underline{A}_{g'}$, then g and g' belong to the same double coset, and that every $\underline{A} \in \mathcal{M}(\mathbb{C})$ is obtained in this way, (ii) identifying the canonical model of Sh_K over $\mathcal{K}$ with $S_{\mathcal{K}}$.

1.3.3 The connected components of Sh_K

Recall that $\mathbf{G}' = \mathbf{SU} = \ker(\nu : \mathbf{G} \to \mathbf{T})$. Since $\mathbf{G}'$ is simple and simply connected, strong approximation holds and

$$\mathbf{G}'(\mathbb{A}) = \mathbf{G}'(\mathbb{Q}) G'_\infty K'_f. \tag{1.35}$$

Here $K' = K \cap \mathbf{G}'(\mathbb{A})$, $K'_f = K \cap \mathbf{G}'(\mathbb{A}_f)$. From the connectedness of G'_∞, we deduce that

$$\mathbf{G}'(\mathbb{Q})\backslash\mathbf{G}'(\mathbb{A})/K' \tag{1.36}$$

is connected.

As $N \geq 3$, $\nu(K) \cap \mathcal{K}^\times = \{1\}$. Here $\mathcal{K}^\times = \nu(\mathbf{G}(\mathbb{Q}))$, and it follows that

$$\mathbf{G}'(\mathbb{Q})\backslash\mathbf{G}'(\mathbb{A})/K' \hookrightarrow \mathbf{G}(\mathbb{Q})\backslash\mathbf{G}(\mathbb{A})/K \tag{1.37}$$

is injective. We now claim (see also Theorem 2.4 and 2.5 of [7]) that

$$\nu : \pi_0(\mathbf{G}(\mathbb{Q})\backslash\mathbf{G}(\mathbb{A})/K) \simeq \pi_0(\mathbf{T}(\mathbb{Q})\backslash\mathbf{T}(\mathbb{A})/\nu(K)) \tag{1.38}$$

is a bijection. For ν is surjective ([7] (0.2)) and continuous (on double coset spaces) so clearly induces a surjective map between the sets of connected components. On the other hand, if $[g_1]$ and $[g_2]$ (double cosets of $g_i \in \mathbf{G}(\mathbb{A})$) are mapped by ν to the same connected component in $\mathbf{T}(\mathbb{Q})\backslash\mathbf{T}(\mathbb{A})/\nu(K)$, then since G_∞ is mapped *onto* the connected component of the identity in $\mathbf{T}(\mathbb{Q})\backslash\mathbf{T}(\mathbb{A})/\nu(K)$, modifying g_1 by an element of G_∞ we may assume that

$$\nu([g_1]) = \nu([g_2]) \in \mathbf{T}(\mathbb{Q})\backslash\mathbf{T}(\mathbb{A})/\nu(K), \tag{1.39}$$

without changing the connected component in which $[g_1]$ lies. Once this has been established, for appropriate representatives g_i of the double cosets, $g_1^{-1}g_2 \in \mathbf{G}'(\mathbb{A})$, so by the connectedness of $\mathbf{G}'(\mathbb{Q})\backslash\mathbf{G}'(\mathbb{A})/K'$, $[g_1]$ and $[g_2]$ lie in the same connected component of $\mathbf{G}(\mathbb{Q})\backslash\mathbf{G}(\mathbb{A})/K$.

The group $\pi_0(\mathbf{T}(\mathbb{Q})\backslash\mathbf{T}(\mathbb{A})/\nu(K))$ is the group

$$\mathcal{K}^\times\backslash\mathcal{K}^\times_{\mathbb{A}}/\mathbb{C}^\times\nu(K_f) = \mathcal{K}^\times\backslash\mathcal{K}^\times_f/\nu(K_f). \tag{1.40}$$

It sits in a short exact sequence

$$0 \to \mu_{\mathcal{K}}\backslash U_{\mathcal{K}}/\nu(K_f) \to \mathcal{K}^\times\backslash\mathcal{K}^\times_f/\nu(K_f) \stackrel{cl}{\to} Cl_{\mathcal{K}} \to 0, \tag{1.41}$$

where $U_{\mathcal{K}}$ is the product of local units at all the finite primes and $Cl_{\mathcal{K}}$ is the class group.

1.3.4 The cl and ν_N invariants of a connected component

The norm $N : \mathcal{K}^\times \to \mathbb{Q}^\times$ satisfies $N \circ \nu = \nu\nu^\rho = \mu$ and hence induces a map

$$\mathcal{K}^\times\backslash\mathcal{K}^\times_f/\nu(K_f) \to \mathbb{Q}^\times_+\backslash\mathbb{Q}^\times_f/\mu(K_f). \tag{1.42}$$

Using the lattice L as an integral structure in V, we see that $\mathbf{G}$ comes from a group scheme $\mathbf{G}_{\mathbb{Z}}$ over $\mathbb{Z}$, whose points in any ring A are

$$\mathbf{G}_{\mathbb{Z}}(A) = \left\{(g, \mu) \in GL_{\mathcal{O}_{\mathcal{K}}\otimes A}(L_A) \times A^\times |\ \langle gu, gv\rangle = \mu\,\langle u, v\rangle\right\}. \tag{1.43}$$

We likewise get that μ is a homomorphism from $\mathbf{G}_{\mathbb{Z}}$ to $\mathbb{G}_m$. The diagram

$$\begin{array}{ll} \mathbf{G}_{\mathbb{Z}}(\mathbb{Z}_p) & \stackrel{\mu}{\to} \mathbb{Z}_p^\times \\ \downarrow & \ \ \downarrow \\ \mathbf{G}_{\mathbb{Z}}(\mathbb{Z}_p/N\mathbb{Z}_p) & \stackrel{\mu}{\to} (\mathbb{Z}_p/N\mathbb{Z}_p)^\times \end{array} \tag{1.44}$$

commutes, $\mathbf{G}_{\mathbb{Z}}(\mathbb{Z}_p) = K_p^0$ and the kernel of $\mathbf{G}_{\mathbb{Z}}(\mathbb{Z}_p) \to \mathbf{G}_{\mathbb{Z}}(\mathbb{Z}_p/N\mathbb{Z}_p)$ is K_p. This shows that $\mu(K_f) \subset \hat{\mathbb{Z}}^\times(N)$, the product of local units congruent to $1 \mod N$. But

$$\mathbb{Q}^\times_+\backslash\mathbb{Q}^\times_f/\hat{\mathbb{Z}}^\times(N) = (\mathbb{Z}/N\mathbb{Z})^\times. \tag{1.45}$$

To conclude, we have shown the existence of two maps from the set of connected components:

$$cl : \pi_0(\mathbf{G}(\mathbb{Q})\backslash\mathbf{G}(\mathbb{A})/K) \to Cl_{\mathcal{K}} \tag{1.46}$$

$$\nu_N : \pi_0(\mathbf{G}(\mathbb{Q})\backslash\mathbf{G}(\mathbb{A})/K) \to (\mathbb{Z}/N\mathbb{Z})^\times. \tag{1.47}$$

These two maps are independent: together they map $\pi_0(\mathbf{G}(\mathbb{Q})\backslash\mathbf{G}(\mathbb{A})/K)$ *onto* $Cl_{\mathcal{K}} \times (\mathbb{Z}/N\mathbb{Z})^\times$. On the other hand, they have a non-trivial common kernel, which grows with N, as is evident from the interpretation of $\mathcal{K}^\times\backslash\mathcal{K}_f^\times/\nu(K_f)$ as the Galois group of a certain class field extension of $\mathcal{K}$. The map cl gives the restriction to the Hilbert class field, while the map ν_N gives the restriction to the cyclotomic field $\mathbb{Q}(\mu_N)$. We have singled out cl and ν_N, because when $N \geq 3$, they have an interpretation in terms of the complex points of Sh_K.

Proposition 1.5 *Let* $[g] \in \mathbf{G}(\mathbb{Q})\backslash\mathbf{G}(\mathbb{A})/K = Sh_K(\mathbb{C})$*. Then*

(i) $cl([g])$ *is the Steinitz class of the lattice* $L_g = g_f(\hat{L}) \cap V$ *in* $Cl_{\mathcal{K}}$*.*
(ii) $\nu_N([g])$ *is (essentially) the* $\nu_{N,g}$ *that appears in the definition of* α_g *(see 1.3.1).*

Proof (i) $cl([g])$ is the class of the ideal $(\nu(g_f))$ associated with the idele $\nu(g_f) \in \mathcal{K}_f^\times$. This ideal is in the same class as $(\det(g_f))$, because $\mu(g_f) \in \mathbb{Q}_f^\times$, so $(\mu(g_f))$ is principal. But the class of $(\det(g_f))$ is the Steinitz class of L_g, since the Steinitz class of L is trivial.

(ii) To find $\nu_N([g])$, we first project the idele $\mu(g_f)$ to $\hat{\mathbb{Z}}^\times$ using $\mathbb{Q}_f^\times = \mathbb{Q}_+^\times\hat{\mathbb{Z}}^\times$. But this is just $\tilde{\mu}(g_f)^{-1}\mu(g_f)$. We then take the result modulo N, so

$$\nu_N([g]) = \tilde{\mu}(g_f)^{-1}\mu(g_f) \mod N. \tag{1.48}$$

Now the definition of the tuple $(A_g, \lambda_g, \iota_g, \alpha_g)$ is such that if $u, v \in N^{-1}L/L$, then

$$\begin{aligned}\langle\alpha_g(u), \alpha_g(v)\rangle_{\lambda_g} &= \exp\left(2\pi i N \langle g_f u, g_f v\rangle_g\right)\\ &= \exp\left(2\pi i \tilde{\mu}(g_f)^{-1} N \langle g_f u, g_f v\rangle\right)\\ &= \exp\left(2\pi i \tilde{\mu}(g_f)^{-1}\mu(g_f) N \langle u, v\rangle\right)\\ &= \exp\left(2\pi i \nu_N([g]) N \langle u, v\rangle\right)\end{aligned} \tag{1.49}$$

Part (ii) follows if we identify $\nu_{N,g} \in Isom_{\mathbb{C}}(N^{-1}\mathbb{Z}/\mathbb{Z}, \mu_N)$ with $\nu_N([g]) \in (\mathbb{Z}/N\mathbb{Z})^\times$ using $\exp(2\pi i(\cdot))$.

1.3.5 The complex uniformization

Recall that $\mathfrak{X} = G_\infty/K_\infty$ and that it was equipped with a base point x_0 (corresponding to $(z, u) = (\delta_{\mathcal{K}}/2, 0)$ in the Siegel domain of the second kind). Let $1 = g_1, \ldots, g_m \in \mathbf{G}(\mathbb{A}_f)$ $(m = \#(\mathcal{K}^\times\backslash\mathcal{K}_f^\times/\nu(K_f)))$ be representatives of the connected components of $\mathbf{G}(\mathbb{Q})\backslash\mathbf{G}(\mathbb{A})/K$, and define congruence groups

$$\Gamma_j = \mathbf{G}(\mathbb{Q}) \cap g_j K_f g_j^{-1}. \tag{1.50}$$

We write $[x, g_j]$ for $\mathbf{G}(\mathbb{Q})(x, g_j K_f) \in \mathbf{G}(\mathbb{Q})\backslash(\mathfrak{X} \times \mathbf{G}(\mathbb{A}_f)/K_f) = \mathbf{G}(\mathbb{Q})\backslash\mathbf{G}(\mathbb{A})/K$. Then $[x', g_j] = [x, g_j]$ if and only if $x' = \gamma x$ for $\gamma \in \Gamma_j$. The map

$$\coprod_{j=1}^m X_{\Gamma_j} = \coprod_{j=1}^m \Gamma_j\backslash\mathfrak{X} \simeq \mathrm{Sh}_K(\mathbb{C}) \tag{1.51}$$

sending $\Gamma_j x$ to $[x, g_j]$ is an isomorphism.

Note that $\Gamma_1 = \Gamma$ is the principal level-N congruence subgroup in $\mathbf{G}_{\mathbb{Z}}(\mathbb{Z})$, the stabilizer of L. Similarly, Γ_j is the principal level-N congruence subgroup in the stabilizer of L_{g_j}, and is thus a group of the type considered in 1.2.1, except that we have dropped the assumption on the Steinitz class of L_{g_j}. As $N \geq 3$, $\det(\gamma) = 1$ and $\mu(\gamma) = 1$ for all $\gamma \in \Gamma_j$, for every j. Indeed, on the one hand these are in $\mathcal{K}^\times$ and $\mathbb{Q}_+^\times$, respectively. On the other hand, they are local units which are congruent to $1 \mod N$ everywhere. It follows that Γ_j are subgroups of $\mathbf{G}'(\mathbb{Q}) = \mathbf{SU}(\mathbb{Q})$.

We get a similar decomposition to connected components (as an algebraic surface)

$$S_{\mathbb{C}} = \coprod_{j=1}^{m} S_{\Gamma_j} \tag{1.52}$$

and we write $S_{\mathbb{C}}^* = \coprod_{j=1}^{m} S_{\Gamma_j}^*$ for the Baily–Borel compactification.

1.4 Smooth compactifications

1.4.1 The smooth compactification of X_Γ

We begin by working in the complex analytic category and follow the exposition of [5]. The Baily–Borel compactification X_Γ^* is singular at the cusps and does not admit a modular interpretation. For general unitary Shimura varieties, the theory of toroidal compactifications provides smooth compactifications that depend, in general, on extra data. It is a unique feature of Picard modular surfaces, stemming from the finiteness of $\mathcal{O}_{\mathcal{K}}^\times$, that this smooth compactification is canonical. As all cusps are equivalent (if we vary the lattice L or Γ), it is enough, as usual, to study the smooth compactification at c_∞. In [5] this is described for an arbitrary L (not even $\mathcal{O}_{\mathcal{K}}$-free), but for simplicity we write it down only for $L = L_0$.

As $N \geq 3$, elements of Γ stabilizing c_∞ lie in N_∞.[2] The computations, which we omit, are somewhat simpler if N is *even*, an assumption made for the rest of this section. Let

$$\Gamma_{\text{cusp}} = \Gamma \cap N_\infty. \tag{1.53}$$

Lemma 1.6 *Let $N \geq 3$ be even. The matrix $n(s,r) \in \Gamma_{\text{cusp}}$ if and only if:* (i) *($d_{\mathcal{K}} \equiv 1 \mod 4$) $s \in N\mathcal{O}_{\mathcal{K}}$, $r \in ND_{\mathcal{K}}\mathbb{Z}$,* (ii) *($d_{\mathcal{K}} \equiv 2, 3 \mod 4$) $s \in N\mathcal{O}_{\mathcal{K}}$ and $r \in 2^{-1}ND_{\mathcal{K}}\mathbb{Z}$.*

Let $M = N|D_{\mathcal{K}}|$ in case (i) and $M = 2^{-1}N|D_{\mathcal{K}}|$ in case (ii). This is the *width of the cusp* c_∞. Let

$$q = q(z) = e^{2\pi i z/M}. \tag{1.54}$$

For $R > 0$, the domain $\Omega_R = \left\{(z,u) \in \mathfrak{X} | \, \lambda(z,u) > R\right\}$ is invariant under Γ_{cusp} and if R is large enough, two points of it are Γ-equivalent if and only if they are Γ_{cusp}-equivalent. A sufficiently small punctured neighborhood of c_∞ in X_Γ^* therefore looks like $\Gamma_{\text{cusp}} \backslash \Omega_R$. As

$$n(s,r)(z,u) = (z + \delta\bar{s}(u + s/2) + r, u + s) \tag{1.55}$$

we obtain the following description of $\Gamma_{\text{cusp}} \backslash \Omega_R$. Let $\Lambda = N\mathcal{O}_{\mathcal{K}}$ and $E = \mathbb{C}/\Lambda$, an elliptic curve with complex multiplication by $\mathcal{O}_{\mathcal{K}}$. Let $\mathcal{T}$ be the quotient

$$\mathcal{T} = (\mathbb{C} \times \mathbb{C})/\Lambda \tag{1.56}$$

where the action of $s \in \Lambda$ is via

$$[s] : (t,u) \mapsto \left(e^{2\pi i \delta\bar{s}(u+s/2)/M} t, u + s\right). \tag{1.57}$$

[2] No confusion should arise from the use of the letter N to denote both the level and the unipotent radical of P.

It is a line bundle over E via the second projection. We denote the class of (t, u) modulo the action of Λ by $[t, u]$.

Proposition 1.7 *Let $\mathcal{T}_R \subset \mathcal{T}$ be the disk bundle consisting of all the points $[t, u]$ where*

$$|t| < e^{-\pi|\delta|(R+u\bar{u})/M}. \tag{1.58}$$

(This condition is invariant under the action of Λ.) Let $\mathcal{T}'_R$ be the punctured disk bundle obtained by removing the zero section from $\mathcal{T}_R$. Then the map $(z, u) \mapsto (q(z), u)$ induces an analytic isomorphism between $\Gamma_{\mathrm{cusp}}\backslash\Omega_R$ and $\mathcal{T}'_R$.

Proof This follows from the discussion so far and the fact that $\lambda(z, u) > R$ is equivalent to the above condition on $t = q(z)$ ([5], Prop. 2.1).

To get a smooth compactification $\bar{X}_\Gamma$ of X_Γ (as a complex surface), we glue the disk bundle $\mathcal{T}_R$ to X_Γ along $\mathcal{T}'_R$. In other words, we complete $\mathcal{T}'_R$ by adding the zero section, which is isomorphic to E. The same procedure should be carried out at any other cusp of $\mathcal{C}_\Gamma$.

Note that the geodesic (1.15) connecting $(z, u) \in \mathfrak{X}$ to the cusp c_∞ projects in $\bar{X}_\Gamma$ to a geodesic which meets E transversally at the point $u \mod \Lambda$. We caution that this geodesic in X_Γ depends on (z, u) and c_∞ and not only on their images modulo Γ.

The line bundle $\mathcal{T}$ is the *inverse* of an ample line bundle on E. In fact, $\mathcal{T}^\vee$ is the N-th (resp. $2N$-th) power of one of the four basic *theta line bundles* if $d_\mathcal{K} \equiv 1 \mod 4$ (resp. $d_\mathcal{K} \equiv 2, 3 \mod 4$). A basic theta function of the lattice Λ satisfies, for $u \in \mathbb{C}$ and $s \in \Lambda$,

$$\theta(u+s) = \alpha(s)e^{2\pi\bar{s}(u+s/2)/|\delta|N^2}\theta(u) \tag{1.59}$$

where $\alpha : \Lambda \to \pm 1$ is a quasi-character (see [31], p. 25). Recalling the relation between M and N, and the assumption that N was even, we easily get the relation between $\mathcal{T}$ and the theta line bundles.

Recall that with any $x = (z, u) \in \mathfrak{X}$ we associated a complex abelian variety A_x, and another model A'_x of the same abelian variety (1.27). This allowed us to define sections $\mathrm{d}\zeta_1$, $\mathrm{d}\zeta_2$ and $\mathrm{d}\zeta_3$ of $\omega_{\mathcal{A}/\mathfrak{X}}$. A simple matrix computation gives the following.

Lemma 1.8 *The sections $d\zeta_1$ and $d\zeta_3$ are invariant under Γ_{cusp}. The section $d\zeta_2$ is invariant modulo the sub-bundle generated by $d\zeta_1$.*

Thus $\mathrm{d}\zeta_1$, $\mathrm{d}\zeta_3$ and $\mathrm{d}\zeta_2 \mod \langle \mathrm{d}\zeta_1\rangle$ descend to well-defined sections in the neighborhood $\mathcal{T}_R \simeq \Gamma_{\mathrm{cusp}}\backslash\Omega_R \cup E$ of E in $\bar{X}_\Gamma$.

1.4.2 The smooth compactification of S

The arithmetic compactification $\bar{S}$ of the Picard surface S (over R_0) is due to Larsen [28,29] (see also [2,26]). We summarize the results in the following theorem. We mention first that as $S_\mathbb{C}$ has a canonical model S over R_0, its Baily–Borel compactification $S^*_\mathbb{C}$ has a similar model S^* over R_0, and S embeds in S^* as an open dense subscheme.

Theorem 1.9 (i) *There exists a projective scheme $\bar{S}$, smooth over R_0, of relative dimension 2, together with an open dense immersion of S in $\bar{S}$, and a proper morphism $p : \bar{S} \to S^*$, making the following diagram commutative*

$$\begin{array}{ccc} S & \rightarrow & \bar{S} \\ \downarrow & \overset{p}{\swarrow} & \\ S^{*} & & \end{array}. \tag{1.60}$$

(ii) *As a complex manifold, there is an isomorphism*

$$\bar{S}_{\mathbb{C}} \simeq \coprod_{j=1}^{m} \bar{X}_{\Gamma_j}, \tag{1.61}$$

extending the isomorphism of $S_{\mathbb{C}}$ *with* $\coprod_{j=1}^{m} X_{\Gamma_j}$.

(iii) *Let* $C = p^{-1}(S^{*}\backslash S)$. *Let* R_N *be the integral closure of* R_0 *in the ray class field* $\mathcal{K}_N$ *of conductor* N *over* $\mathcal{K}$. *Then the connected components of* C_{R_N} *are geometrically irreducible and are indexed by the cusps of* $S^{*}_{R_N}$ *over which they sit. Furthermore, each component* $E \subset C_{R_N}$ *is an elliptic curve with complex multiplication by* $\mathcal{O}_{\mathcal{K}}$.

We call C the *cuspidal divisor*. If $c \in S^{*}_{\mathbb{C}}\backslash S_{\mathbb{C}}$ is a cusp, we denote the complex elliptic curve $p^{-1}(c)$ by E_c. Although E_c is in principle definable over the Hilbert class field $\mathcal{K}_1$, no canonical model of it over that field is provided by $\bar{S}$. On the other hand, E_c does come with a canonical model over $\mathcal{K}_N$, and even over R_N.

We refer to [2,28] for a moduli-theoretic interpretation of C as a moduli space for semi-abelian schemes with a suitable action of $\mathcal{O}_{\mathcal{K}}$ and a "level-N structure".

1.4.3 Change of level

Assume that $N \geq 3$ is even, and $N' = QN$. We then obtain a covering map $X_{\Gamma(N')} \rightarrow X_{\Gamma(N)}$ where by $\Gamma(N)$ we denote the group previously denoted by Γ. Near any of the cusps, the analytic model allows us to analyze this map locally. Let E' be an irreducible cuspidal component of $\bar{X}_{\Gamma(N')}$ mapping to the irreducible component E of $\bar{X}_{\Gamma(N)}$. The following is a consequence of the discussion in the previous sections.

Proposition 1.10 *The map* $E' \rightarrow E$ *is a multiplication-by-*Q *isogeny, hence étale of degree* Q^2. *When restricted to a neighborhood of* E', *the covering* $\bar{X}_{\Gamma(N')} \rightarrow \bar{X}_{\Gamma(N)}$ *is of degree* Q^3 *and has ramification index* Q *along* E, *in the normal direction to* E.

Corollary 1.11 *The pull-back to* E' *of the normal bundle* $\mathcal{T}(N)$ *of* E *is the* Q*th power of the normal bundle* $\mathcal{T}(N')$ *of* E'.

1.5 The universal semi-abelian scheme $\mathcal{A}$

1.5.1 The universal semi-abelian scheme over $\bar{S}$

As Larsen and Bellaïche explain, the universal abelian scheme $\pi : \mathcal{A} \rightarrow S$ extends canonically to a semi-abelian scheme $\pi : \mathcal{A} \rightarrow \bar{S}$. The polarization λ extends over the boundary $C = \bar{S}\backslash S$ to a principal polarization λ of the abelian part of $\mathcal{A}$. The action ι of $\mathcal{O}_{\mathcal{K}}$ extends to an action on the semi-abelian variety, which necessarily induces separate actions on the toric part and on the abelian part.

Let E be a connected component of C_{R_N}, mapping (over $\mathbb{C}$ and under the projection p) to the cusp $c \in S^{*}_{\mathbb{C}}$. Then there exist (1) a principally polarized elliptic curve B defined over R_N, with complex multiplication by $\mathcal{O}_{\mathcal{K}}$ and CM type Σ, and (2) an ideal $\mathfrak{a}$ of $\mathcal{O}_{\mathcal{K}}$, such that every fiber $\mathcal{A}_x$ of $\mathcal{A}$ over E is an $\mathcal{O}_{\mathcal{K}}$-group extension of B by the $\mathcal{O}_{\mathcal{K}}$-torus $\mathfrak{a} \otimes \mathbb{G}_m$. Both B (with its polarization) and the ideal class $[\mathfrak{a}] \in Cl_{\mathcal{K}}$ are uniquely determined by the cusp c. Only the extension class in the category of $\mathcal{O}_{\mathcal{K}}$-groups varies as we move along E. Note that since the Lie algebra of the torus is of type (1, 1), the Lie algebra of such an extension $\mathcal{A}_x$ is of type (2, 1), as is the case at an interior point $x \in S$. If we extend scalars to $\mathbb{C}$, the isomorphism type of B is given by another ideal class $[\mathfrak{b}]$ (i.e. $B(\mathbb{C}) \simeq \mathbb{C}/\mathfrak{b}$). In this case, we say that the cusp c is of type $(\mathfrak{a}, \mathfrak{b})$.

The above discussion defines a homomorphism (of fppf sheaves over $\mathrm{Spec}(R_N)$)

$$E \to Ext^1_{\mathcal{O}_{\mathcal{K}}}(B, \mathfrak{a} \otimes \mathbb{G}_m). \tag{1.62}$$

As we shall see soon, the *Ext* group is represented by an elliptic curve with CM by $\mathcal{O}_{\mathcal{K}}$, defined over R_N, and this map is an isogeny.

1.5.2 $\mathcal{O}_{\mathcal{K}}$-semi-abelian schemes of type $(\mathfrak{a}, \mathfrak{b})$

We digress to discuss the moduli space for semi-abelian schemes of the type found above points of E. Let R be an R_0-algebra, B an elliptic curve over R with complex multiplication by $\mathcal{O}_{\mathcal{K}}$ and CM type Σ, and $\mathfrak{a}$ an ideal of $\mathcal{O}_{\mathcal{K}}$. Consider a semi-abelian scheme $\mathcal{G}$ over R, endowed with an $\mathcal{O}_{\mathcal{K}}$ action $\iota : \mathcal{O}_{\mathcal{K}} \to End(\mathcal{G})$, and a short exact sequence

$$0 \to \mathfrak{a} \otimes \mathbb{G}_m \to \mathcal{G} \to B \to 0 \tag{1.63}$$

of $\mathcal{O}_{\mathcal{K}}$-group schemes over R. We call all this data a *semi-abelian scheme of type* $(\mathfrak{a}, B)$ (over R). The group classifying such structures is $Ext^1_{\mathcal{O}_{\mathcal{K}}}(B, \mathfrak{a} \otimes \mathbb{G}_m)$. Any $\chi \in \mathfrak{a}^* = \mathrm{Hom}(\mathfrak{a}, \mathbb{Z})$ defines, by push-out, an extension $\mathcal{G}_\chi$ of B by $\mathbb{G}_m$, hence a point of $B^t = Ext^1(B, \mathbb{G}_m)$. We therefore get a homomorphism from $Ext^1_{\mathcal{O}_{\mathcal{K}}}(B, \mathfrak{a} \otimes \mathbb{G}_m)$ to $\mathrm{Hom}(\mathfrak{a}^*, B^t)$. A simple check shows that its image is in $\mathrm{Hom}_{\mathcal{O}_{\mathcal{K}}}(\mathfrak{a}^*, B^t) = \delta_{\mathcal{K}}\mathfrak{a} \otimes_{\mathcal{O}_{\mathcal{K}}} B^t$ and that this construction yields an isomorphism

$$\mathrm{Ext}^1_{\mathcal{O}_{\mathcal{K}}}(B, \mathfrak{a} \otimes \mathbb{G}_m) \simeq \delta_{\mathcal{K}}\mathfrak{a} \otimes_{\mathcal{O}_{\mathcal{K}}} B^t. \tag{1.64}$$

Here we have used the canonical identification $\mathfrak{a}^* = \delta_{\mathcal{K}}^{-1}\mathfrak{a}^{-1}$ (via the trace pairing). Although $(\delta_{\mathcal{K}})$ is a principal ideal, so can be ignored, it is better to keep track of its presence. We emphasize that the CM type of B^t, with the natural action of $\mathcal{O}_{\mathcal{K}}$ derived from its action on B, is $\bar{\Sigma}$ rather than Σ.

Thus over $\delta_{\mathcal{K}}\mathfrak{a} \otimes_{\mathcal{O}_{\mathcal{K}}} B^t$, there is a universal semi-abelian scheme $\mathcal{G}(\mathfrak{a}, B)$ of type $(\mathfrak{a}, B)$, and any $\mathcal{G}$ as above, over any base R'/R, is obtained from $\mathcal{G}(\mathfrak{a}, B)$ by pull-back (specialization) with respect to a unique map $\mathrm{Spec}(R') \to \delta_{\mathcal{K}}\mathfrak{a} \otimes_{\mathcal{O}_{\mathcal{K}}} B^t$.

When $R = \mathbb{C}$, $B \simeq \mathbb{C}/\mathfrak{b}$ for a unique ideal class $[\mathfrak{b}]$ (with $\mathcal{O}_{\mathcal{K}}$ acting via Σ). Then, canonically, $B^t = \mathbb{C}/\delta_{\mathcal{K}}^{-1}\overline{\mathfrak{b}}^{-1}$ (with $\mathcal{O}_{\mathcal{K}}$ acting via $\bar{\Sigma}$). The pairing between the lattices, $\mathfrak{b} \times \delta_{\mathcal{K}}^{-1}\overline{\mathfrak{b}}^{-1} \to \mathbb{Z}$ is $(x, y) \mapsto Tr_{\mathcal{K}/\mathbb{Q}}(x\bar{y})$. Since the $\mathcal{O}_{\mathcal{K}}$ action on B^t is via $\bar{\Sigma}$,

$$\mathrm{Ext}^1_{\mathcal{O}_{\mathcal{K}}}(\mathbb{C}/\mathfrak{b}, \mathfrak{a} \otimes \mathbb{G}_m) \simeq \delta_{\mathcal{K}}\mathfrak{a} \otimes_{\mathcal{O}_{\mathcal{K}}} \mathbb{C}/\delta_{\mathcal{K}}^{-1}\overline{\mathfrak{b}}^{-1} = \mathbb{C}/\bar{\mathfrak{a}}\overline{\mathfrak{b}}^{-1}. \tag{1.65}$$

The universal semi-abelian variety $\mathcal{G}(\mathfrak{a}, B)$ will now be denoted $\mathcal{G}(\mathfrak{a}, \mathfrak{b})$. In 1.6.2, we give a complex analytic model of this $\mathcal{G}(\mathfrak{a}, \mathfrak{b})$.

1.6 Degeneration of $\mathcal{A}$ along a geodesic connecting to a cusp

1.6.1 The degeneration to a semi-abelian variety

It is instructive to use the "moving lattice model" to compute the degeneration of the universal abelian scheme along a geodesic, as we approach a cusp. To simplify the computations, assume for the rest of this section, as before, that $N \geq 3$ is *even* and that the cusp is the standard cusp at infinity $c = c_\infty$. In this case, we have shown that $E_c = \mathbb{C}/\Lambda$, where $\Lambda = N\mathcal{O}_{\mathcal{K}}$, and we have given a neighborhood of E_c in $\bar{X}_\Gamma$ the structure of a disk bundle in a line bundle $\mathcal{T}$. See Proposition 1.7.

Consider the geodesic (1.15) connecting (z, u) to c_∞. Consider the universal abelian scheme in the moving lattice model [*cf* (1.27)]. Of the three vectors used to span L'_x over $\mathcal{O}_{\mathcal{K}}$ in (1.25), the first two do not depend on z. As u is fixed along the geodesic, they are not changed. The third vector represents a cycle that vanishes at the cusp (together with all its $\mathcal{O}_{\mathcal{K}}$-multiples). We conclude that A'_x degenerates to

$$\mathbb{C}^3/\mathrm{Span}_{\iota'(\mathcal{O}_{\mathcal{K}})}\left\{\begin{pmatrix}0\\1\\1\end{pmatrix},\begin{pmatrix}1\\0\\u\end{pmatrix}\right\}. \tag{1.66}$$

Making the change of variables $(\zeta'_1, \zeta'_2, \zeta'_3) = (\zeta_1, \zeta_2 + \bar{u}\zeta_1, \zeta_3)$ does not alter the $\mathcal{O}_{\mathcal{K}}$ action and gives the more symmetric model

$$\mathcal{G}_u = \mathbb{C}^3/\mathrm{Span}_{\iota'(\mathcal{O}_{\mathcal{K}})}\left\{\begin{pmatrix}0\\1\\1\end{pmatrix},\begin{pmatrix}1\\\bar{u}\\u\end{pmatrix}\right\} \tag{1.67}$$

(but note that ζ'_2, unlike ζ_2, does not vary holomorphically in the *family* $\{\mathcal{G}_u\}$, only in each fiber individually).

Let $e(x) = e^{2\pi i x} : \mathbb{C} \to \mathbb{C}^\times$ be the exponential map, with kernel $\mathbb{Z}$. For any ideal $\mathfrak{a}$ of $\mathcal{O}_{\mathcal{K}}$, it induces a map

$$e_{\mathfrak{a}} : \mathfrak{a} \otimes \mathbb{C} \to \mathfrak{a} \otimes \mathbb{C}^\times \tag{1.68}$$

with kernel $\mathfrak{a} \otimes 1$. As usual we identify $\mathfrak{a} \otimes \mathbb{C}$ with $\mathbb{C}(\Sigma) \oplus \mathbb{C}(\bar{\Sigma})$, sending $a \otimes \zeta \mapsto (a\zeta, \bar{a}\zeta)$. We now note that if we use this identification to identify $\mathbb{C}^3$ with $\mathbb{C} \oplus (\mathcal{O}_{\mathcal{K}} \otimes \mathbb{C})$ (an identification which is compatible with the $\mathcal{O}_{\mathcal{K}}$ action), then the $\iota'(\mathcal{O}_{\mathcal{K}})$-span of the vector ${}^t(0, 1, 1)$ is just the kernel of $e_{\mathcal{O}_{\mathcal{K}}}$. We conclude that

$$\mathcal{G}_u \simeq \{\mathbb{C} \oplus (\mathcal{O}_{\mathcal{K}} \otimes \mathbb{C}^\times)\}/L_u \tag{1.69}$$

where L_u is the sub-$\mathcal{O}_{\mathcal{K}}$-module

$$L_u = \left\{(s, e_{\mathcal{O}_{\mathcal{K}}}(s\bar{u}, \bar{s}u)) | \, s \in \mathcal{O}_{\mathcal{K}}\right\}. \tag{1.70}$$

This clearly gives $\mathcal{G}_u$ the structure of an $\mathcal{O}_{\mathcal{K}}$-semi-abelian variety of type $(\mathcal{O}_{\mathcal{K}}, \mathcal{O}_{\mathcal{K}})$, i.e. an extension

$$0 \to \mathcal{O}_{\mathcal{K}} \otimes \mathbb{C}^\times \to \mathcal{G}_u \to \mathbb{C}/\mathcal{O}_{\mathcal{K}} \to 0. \tag{1.71}$$

1.6.2 The analytic uniformization of the universal semi-abelian variety of type $(\mathfrak{a}, \mathfrak{b})$

We now compare the description that we have found for the degeneration of $\mathcal{A}$ along the geodesic connecting (z, u) to c_∞ with the analytic description of the universal semi-abelian variety of type $(\mathfrak{a}, \mathfrak{b})$.

Proposition 1.12 *Let $\mathfrak{a}$ and $\mathfrak{b}$ be two ideals of $\mathcal{O}_{\mathcal{K}}$. For $u \in \mathbb{C}$ consider*

$$\mathcal{G}_u \simeq \{\mathbb{C} \oplus (\mathfrak{a} \otimes \mathbb{C}^\times)\}/L_u \tag{1.72}$$

where

$$L_u = \left\{(s, e_{\mathfrak{a}}(s\bar{u}, \bar{s}u)) | \, s \in \mathfrak{b}\right\}. \tag{1.73}$$

Then $\mathcal{G}_u$ is a semi-abelian variety of type $(\mathfrak{a}, \mathfrak{b})$, any complex semi-abelian variety of this type is a $\mathcal{G}_u$, and $\mathcal{G}_u \simeq \mathcal{G}_v$ if and only if $u - v \in \overline{\mathfrak{a}\mathfrak{b}^{-1}}$.

Proof That $\mathcal{G}_u$ is a semi-abelian variety of type $(\mathfrak{a}, \mathfrak{b})$ is obvious. That any abelian variety of this type is a $\mathcal{G}_u$ follows by passing to the universal cover $\mathbb{C}^2(\Sigma) \oplus \mathbb{C}(\bar{\Sigma})$, and noting that by a change of variables in the Σ- and $\bar{\Sigma}$-isotypical parts, we may assume that the lattice by which we divide is of the form

$$\mathfrak{a}\begin{pmatrix}0\\1\\1\end{pmatrix} \oplus \mathfrak{b}\begin{pmatrix}1\\\bar{u}\\u\end{pmatrix}. \tag{1.74}$$

Finally, the map $u \mapsto [\mathcal{G}_u]$ is a homomorphism $\mathbb{C} \to \mathit{Ext}^1_{\mathcal{O}_\mathcal{K}}(\mathbb{C}/\mathfrak{b}, \mathfrak{a} \otimes \mathbb{C}^\times)$, so we only have to prove that $\mathcal{G}_u$ is split if and only if $u \in \bar{\mathfrak{a}}\overline{\mathfrak{b}}^{-1}$. But one can check easily that $\mathcal{G}_u$ is trivial if and only if $(s\bar{u}, \bar{s}u) \in \ker e_\mathfrak{a} = \mathfrak{a} \otimes 1 = \{(a, \bar{a}) | a \in \mathfrak{a}\}$ for every $s \in \mathfrak{b}$, and this holds if and only if $u \in \bar{\mathfrak{a}}\overline{\mathfrak{b}}^{-1}$.

Corollary 1.13 *Let $N \geq 3$ be even. Let $c = c_\infty$ be the cusp at infinity. Then the map*

$$E_c \to \mathit{Ext}^1_{\mathcal{O}_\mathcal{K}}(\mathbb{C}/\mathcal{O}_\mathcal{K}, \mathcal{O}_\mathcal{K} \otimes \mathbb{C}^\times) \tag{1.75}$$

sending u to the isomorphism class of the semi-abelian variety above $u \mod \Lambda$ is the isogeny of multiplication by N.

Proof In view of the computations above, and the description of a neighborhood of E_c in $\bar{X}_\Gamma$ given in Proposition 1.7 this map is identified with the canonical map

$$\mathbb{C}/N\mathcal{O}_\mathcal{K} \to \mathbb{C}/\mathcal{O}_\mathcal{K}. \tag{1.76}$$

The extra data carried by $u \in E_c$, which are forgotten by the map of the corollary, come from the level N structure. As mentioned before, according to [28] and [2] the cuspidal divisor C has a modular interpretation as the moduli space for semi-abelian schemes of the type considered above, together with level-N structure ($\mathcal{M}_{\infty,N}$ structures in the language of [2]). A level-N structure on a *semi-abelian* variety $\mathcal{G}$ of type $(\mathfrak{a}, \mathfrak{b})$ consists of (i) a level-N structures $\alpha : N^{-1}\mathcal{O}_\mathcal{K}/\mathcal{O}_\mathcal{K} \simeq \mathfrak{a} \otimes \mu_N$ on the toric part (ii) a level-N structure $\beta : N^{-1}\mathcal{O}_\mathcal{K}/\mathcal{O}_\mathcal{K} \simeq N^{-1}\mathfrak{b}/\mathfrak{b} = B[N]$ on the abelian part (iii) an $\mathcal{O}_\mathcal{K}$-splitting γ of the map $\mathcal{G}[N] \to B[N]$.

Over $c = c_\infty$, when $\mathfrak{a} = \mathfrak{b} = \mathcal{O}_\mathcal{K}$, there are obvious natural choices for α and β (independent of u), but the splittings γ in (iii) form a torsor under $\mathcal{O}_\mathcal{K}/N\mathcal{O}_\mathcal{K}$. If we consider the splitting

$$\gamma_u : N^{-1}\mathcal{O}_\mathcal{K}/\mathcal{O}_\mathcal{K} \ni s \mapsto (s, e_{\mathcal{O}_\mathcal{K}}(s\bar{u}, \bar{s}u)) \mod L_u \tag{1.77}$$

then the tuples $(\mathcal{G}_u, \alpha, \beta, \gamma_u)$ and $(\mathcal{G}_v, \alpha, \beta, \gamma_v)$ are isomorphic if and only if $u \equiv v \mod N\mathcal{O}_\mathcal{K}$, i.e. if and only if u and v represent the same point of E_c.

1.7 The basic automorphic vector bundles

1.7.1 Definition and first properties

Recall that we have denoted by $\pi : \mathcal{A} \to \bar{S}$ the universal semi-abelian variety over $\bar{S}$ (over the base ring R_0). Let $\omega_\mathcal{A}$ be the relative cotangent space at the origin of $\mathcal{A}$. If $e : \bar{S} \to \mathcal{A}$ is the zero section,

$$\omega_\mathcal{A} = e^*\left(\Omega^1_{\mathcal{A}/\bar{S}}\right). \tag{1.78}$$

This is a rank 3 vector bundle over $\bar{S}$ and the action of $\mathcal{O}_\mathcal{K}$ allows to decompose it according to types. We let

$$\mathcal{P} = \omega_\mathcal{A}(\Sigma), \quad \mathcal{L} = \omega_\mathcal{A}(\bar{\Sigma}). \tag{1.79}$$

Then $\mathcal{P}$ is a plane bundle, and $\mathcal{L}$ a line bundle.

Over S (but not over the cuspidal divisor $C = \bar{S}\backslash S$), we have the usual identification $\omega_\mathcal{A} = \pi_*\Omega^1_{\mathcal{A}/S}$. The relative de Rham cohomology of $\mathcal{A}/S$ is a rank 6 vector bundle sitting in an exact sequence (the Hodge filtration)

$$0 \to \omega_\mathcal{A} \to H^1_{dR}(\mathcal{A}/S) \to R^1\pi_*\mathcal{O}_\mathcal{A} \to 0. \tag{1.80}$$

Since, for any abelian scheme, $R^1\pi_*\mathcal{O}_\mathcal{A} = \omega^\vee_{\mathcal{A}^t}$ (canonical isomorphism, see [31]), and $\lambda : \mathcal{A} \to \mathcal{A}^t$ is an isomorphism which reverses CM types, we obtain an exact sequence

$$0 \to \omega_\mathcal{A} \to H^1_{dR}(\mathcal{A}/S) \to \omega^\vee_\mathcal{A}(\rho) \to 0. \tag{1.81}$$

The notation $\mathcal{M}(\rho)$ means that $\mathcal{M}$ is a vector bundle with an $\mathcal{O}_\mathcal{K}$ action and in $\mathcal{M}(\rho)$ the vector bundle structure is that of $\mathcal{M}$, but the $\mathcal{O}_\mathcal{K}$ action is conjugated. Decomposing according to types, we have two short exact sequences

$$\begin{aligned} &0 \to \mathcal{P} \to H^1_{dR}(\mathcal{A}/S)(\Sigma) \to \mathcal{L}^\vee(\rho) \to 0 \\ &0 \to \mathcal{L} \to H^1_{dR}(\mathcal{A}/S)(\bar{\Sigma}) \to \mathcal{P}^\vee(\rho) \to 0. \end{aligned} \tag{1.82}$$

The pairing $\langle,\rangle_\lambda$ on $H^1_{dR}(\mathcal{A}/S)$ induced by the polarization is $\mathcal{O}_S$-linear, alternating, perfect, and satisfies $\langle\iota(a)x, y\rangle_\lambda = \langle x, \iota(\bar{a})y\rangle_\lambda$. It follows that $H^1_{dR}(\mathcal{A}/S)(\Sigma)$ and $H^1_{dR}(\mathcal{A}/S)(\bar{\Sigma})$ are maximal isotropic subspaces and are set in duality. As $\omega_\mathcal{A}$ is also isotropic, this pairing induces pairings

$$\mathcal{P} \times \mathcal{P}^\vee(\rho) \to \mathcal{O}_S, \ \mathcal{L} \times \mathcal{L}^\vee(\rho) \to \mathcal{O}_S. \tag{1.83}$$

These two pairings are the tautological pairings between a vector bundle and its dual.

Another consequence of this discussion that we wish to record is the canonical isomorphism over S

$$\det \mathcal{P} = \mathcal{L}(\rho) \otimes \det\left(H^1_{dR}(\mathcal{A}/S)(\Sigma)\right). \tag{1.84}$$

1.7.2 The factors of automorphy corresponding to $\mathcal{L}$ and $\mathcal{P}$

The formulae below can be deduced also from the matrix calculations in the first few pages of [36]. Let $\Gamma = \Gamma_j$ be one of the groups used in the complex uniformization of $S_\mathbb{C}$, *cf* Sect. 1.3.5. Via the analytic isomorphism $X_\Gamma \simeq S_\Gamma$ with the jth connected component, the vector bundles $\mathcal{P}$ and $\mathcal{L}$ are pulled back to X_Γ and then to the symmetric space $\mathfrak{X}$, where they can be trivialized, hence described by means of factors of automorphy. Let us denote by $\mathcal{P}_{an}$ and $\mathcal{L}_{an}$ the two vector bundles on X_Γ, in the complex analytic category, or their pull-backs to $\mathfrak{X}$.

To trivialize $\mathcal{L}_{an}$, we must choose a nowhere vanishing global section over $\mathfrak{X}$. As usual, we describe it only on the connected component containing the standard cusp, corresponding to $j = 1$ (where $L = L_{g_1} = L_0$). Recalling the "moving lattice model" (1.27) and the coordinates $\zeta_1, \zeta_2, \zeta_3$ introduced there, we note that $\mathrm{d}\zeta_3$ is a generator of $\mathcal{L}_{an} = \omega_\mathcal{A}(\bar{\Sigma})$. For reasons that will become clear later (*cf* Sect. 1.12), we use $2\pi i \cdot \mathrm{d}\zeta_3$ to trivialize $\mathcal{L}_{an}$ over $\mathfrak{X}$. Suppose

$$\gamma = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \in \Gamma \subset SU_\infty. \tag{1.85}$$

If $\gamma(z,u)=(z',u')$, then

$$z'=\frac{a_1z+b_1u+c_1}{a_3z+b_3u+c_3},\quad u'=\frac{a_2z+b_2u+c_2}{a_3z+b_3u+c_3} \tag{1.86}$$

and

$$\gamma\begin{pmatrix}z\\u\\1\end{pmatrix}=j(\gamma;z,u)\begin{pmatrix}z'\\u'\\1\end{pmatrix},\quad j(\gamma;z,u)=a_3z+b_3u+c_3. \tag{1.87}$$

Lemma 1.14 *The following relation holds for every $\gamma\in U_\infty$*

$$\lambda(z,u)=\lambda(\gamma(z,u))\cdot|j(\gamma;z,u)|^2. \tag{1.88}$$

Proof Let $v=v(z,u)={}^t(z,u,1)$. Then

$$\lambda(z,u)=-(v,v). \tag{1.89}$$

As $v(\gamma(z,u))=j(\gamma;z,u)^{-1}\cdot\gamma(v(z,u))$ the lemma follows from $(\gamma v,\gamma v)=(v,v)$.

Let $\mathcal{V}=Lie(\mathcal{A}/\mathfrak{X})=\omega^\vee_{\mathcal{A}/\mathfrak{X}}$ and $\mathcal{W}=\mathcal{V}(\bar{\Sigma})=\mathcal{L}^\vee_{an}$ (a line bundle). At a point $x=(z,u)\in\mathfrak{X}$ the fiber $\mathcal{V}_x$ is identified canonically with $(V_\mathbb{R},J_x)$ and then $\mathcal{W}_x=W_x=\mathbb{C}\cdot{}^t(z,u,1)$.

Proposition 1.15 *For $x=(z,u)\in\mathfrak{X}$ let*

$$v_3(z,u)=\lambda(z,u)^{-1}\begin{pmatrix}z\\u\\1\end{pmatrix}\in\mathcal{W}_x. \tag{1.90}$$

Then (i) *$v_3(z,u)$ is a nowhere vanishing holomorphic section of $\mathcal{W}$,* (ii) $\langle d\zeta_3,v_3\rangle\equiv 1$, (iii) *the automorphy factor corresponding to $d\zeta_3$ is the function $j(\gamma;z,u)$.*

Proof Since, by construction, $\mathrm{d}\zeta_3$ is a nowhere vanishing holomorphic section of $\mathcal{L}$ (over $\mathfrak{X}$), (i) follows from (ii). To prove (ii), we transfer $v_3(z,u)$ to the moving lattice model and get ${}^t(0,0,1)$, which is the dual vector to $\mathrm{d}\zeta_3$. To prove (iii), we compute in $V_\mathbb{R}$ (with the original complex structure!)

$$\frac{\gamma_*v_3(z,u)}{v_3(\gamma(z,u))}=\frac{\lambda(\gamma(z,u))}{\lambda(z,u)}j(\gamma;z,u)=\overline{j(\gamma;z,u)}^{-1}, \tag{1.91}$$

and recall that since $W_{\gamma(z,u)}$ is precisely the line where the complex structure in $(V_\mathbb{R},J_{\gamma(z,u)})$ has been reversed, in $(V_\mathbb{R},J_{\gamma(z,u)})$ we have

$$\frac{\gamma_*v_3(z,u)}{v_3(\gamma(z,u))}=j(\gamma;z,u)^{-1}. \tag{1.92}$$

Dualizing, we get $(x=(z,u))$

$$\frac{(\gamma^{-1})^*\mathrm{d}\zeta_3|_x}{\mathrm{d}\zeta_3|_{\gamma(x)}}=j(\gamma,x). \tag{1.93}$$

This concludes the proof.

Consider next the plane bundle $\mathcal{P}_{an}$. As we will only be interested in scalar-valued modular forms, we do not compute its matrix-valued factor of automorphy (but see [36]). It is important to know, however, that the line bundle $\det\mathcal{P}_{an}$ gives nothing new.

Proposition 1.16 *There is an isomorphism of analytic line bundles over X_Γ,*

$$\det \mathcal{P}_{an} \simeq \mathcal{L}_{an}. \tag{1.94}$$

Moreover, $d\zeta_1 \wedge d\zeta_2$ is a nowhere vanishing holomorphic section of $\det \mathcal{P}_{an}$ *over $\mathfrak{X}$, and the factor of automorphy corresponding to it is $j(\gamma; z, u)$.*

Proof Since a holomorphic line bundle on $X_\Gamma = \Gamma\backslash\mathfrak{X}$ is determined, up to an isomorphism, by its factor of automorphy, and $j(\gamma; z, u)$ is the factor of automorphy of $\mathcal{L}_{an}$ corresponding to $\mathrm{d}\zeta_3$, it is enough to prove the second statement. Let $\mathcal{U} = \mathcal{V}(\Sigma)$ be the plane bundle dual to $\mathcal{P}_{an}$. Let

$$v_1(z,u) = -\lambda(z,u)^{-1}\begin{pmatrix} \bar{u}z \\ (z-\bar{z})/\delta \\ \bar{u} \end{pmatrix} \tag{1.95}$$

and

$$v_2(z,u) = -\lambda(z,u)^{-1}\begin{pmatrix} \bar{z}+\delta u\bar{u} \\ u \\ 1 \end{pmatrix} \tag{1.96}$$

(considered as vectors in $(V_\mathbb{R}, J_x) = \mathcal{V}_x$). As we have seen in (1.27), these two vector fields are sections of $\mathcal{U}$ and at each point $x \in \mathfrak{X}$ form a basis dual to $\mathrm{d}\zeta_1$ and $\mathrm{d}\zeta_2$. It follows that they are holomorphic sections and that $v_1 \wedge v_2$ is the basis dual to $\mathrm{d}\zeta_1 \wedge \mathrm{d}\zeta_2$. We must show that the factor of automorphy corresponding to $v_1 \wedge v_2$ is $j(\gamma; z, u)^{-1}$, i.e. that

$$\frac{\gamma_*(v_1 \wedge v_2(z,u))}{v_1 \wedge v_2(\gamma(z,u))} = j(\gamma; z,u)^{-1}. \tag{1.97}$$

Working in $V_\mathbb{R} = \mathbb{C}^3$ (with the original complex structure)

$$\begin{aligned} &\frac{\gamma_*(v_1 \wedge v_2(z,u))}{v_1 \wedge v_2(\gamma(z,u))} \cdot \frac{1}{\overline{j(\gamma;z,u)}} \\ &\quad = \frac{\gamma_*(v_1 \wedge v_2(z,u))}{v_1 \wedge v_2(\gamma(z,u))} \cdot \frac{\gamma_* v_3(z,u)}{v_3(\gamma(z,u))} = \frac{\gamma_*(v_1 \wedge v_2 \wedge v_3(z,u))}{v_1 \wedge v_2 \wedge v_3(\gamma(z,u))}. \end{aligned} \tag{1.98}$$

But

$$v_1 \wedge v_2 \wedge v_3(z,u) = \delta\lambda(z,u)^{-1} e_1 \wedge e_2 \wedge e_3, \tag{1.99}$$

because

$$\det\begin{pmatrix} \bar{u}z & \bar{z}+\delta u\bar{u} & z \\ (z-\bar{z})/\delta & u & u \\ \bar{u} & 1 & 1 \end{pmatrix} = \delta\lambda(z,u)^2. \tag{1.100}$$

As $\det(\gamma) = 1$, this gives

$$\frac{\gamma_*(v_1 \wedge v_2(z,u))}{v_1 \wedge v_2(\gamma(z,u))} \cdot \frac{1}{\overline{j(\gamma;z,u)}} = \frac{\lambda(\gamma(z,u))}{\lambda(z,u)} = \frac{1}{j(\gamma;z,u)\overline{j(\gamma;z,u)}}, \tag{1.101}$$

and the proof is complete.

1.7.3 The relation $\det \mathcal{P} \simeq \mathcal{L}$ *over* $\bar{S}_\mathcal{K}$

The isomorphism between $\det \mathcal{P}$ and $\mathcal{L}$ is in fact algebraic and even extends to the generic fiber $\bar{S}_\mathcal{K}$ of the smooth compactification.

Proposition 1.17 *One has* $\det\mathcal{P} \simeq \mathcal{L}$ *over* $\bar{S}_{\mathcal{K}}$.

Proof Since $Pic(\bar{S}_{\mathcal{K}}) \subset Pic(\bar{S}_{\mathbb{C}})$, it is enough to prove the proposition over $\mathbb{C}$. By GAGA, it is enough to establish the triviality of $\det\mathcal{P} \otimes \mathcal{L}^{-1}$ in the analytic category. For each connected component X_Γ of $S_{\mathbb{C}}$, the section $(\mathrm{d}\zeta_1 \wedge \mathrm{d}\zeta_2) \otimes \mathrm{d}\zeta_3^{-1}$ descends from $\mathfrak{X}$ to X_Γ, because $\mathrm{d}\zeta_1 \wedge \mathrm{d}\zeta_2$ and $\mathrm{d}\zeta_3$ have the same factor of automorphy $j(\gamma, x)$ $(\gamma \in \Gamma, x \in \mathfrak{X})$. This section is nowhere vanishing on X_Γ and extends to a nowhere vanishing section on $\bar{X}_\Gamma$, trivializing $\det\mathcal{P} \otimes \mathcal{L}^{-1}$. In fact, if c is the standard cusp, $\mathrm{d}\zeta_1 \wedge \mathrm{d}\zeta_2$ and $\mathrm{d}\zeta_3$ are already well defined and nowhere vanishing sections of $\det\mathcal{P}$ and $\mathcal{L}$ in the neighborhood

$$\overline{\Gamma_{\mathrm{cusp}} \backslash \Omega_R} = (\Gamma_{\mathrm{cusp}} \backslash \Omega_R) \cup E_c \tag{1.102}$$

of E_c (see 1.4.1). This is a consequence of the fact that $j(\gamma, x) = 1$ for $\gamma \in \Gamma_{\mathrm{cusp}}$.

An alternative proof is to use Theorem 4.8 of [14]. In our case, it gives a *functor* $\mathcal{V} \mapsto [\mathcal{V}]$ from the category of $\mathbf{G}(\mathbb{C})$-equivariant vector bundles on the compact dual $\mathbb{P}^2_{\mathbb{C}}$ of Sh_K to the category of vector bundles with $\mathbf{G}(\mathbb{A}_f)$-action on the inverse system of Shimura varieties Sh_K. Here $\mathbb{P}^2_{\mathbb{C}} = \mathbf{G}(\mathbb{C})/\mathbf{H}(\mathbb{C})$, where $\mathbf{H}(\mathbb{C})$ is the parabolic group stabilizing the line $\mathbb{C} \cdot {}^t(\delta/2, 0, 1)$ in $\mathbf{G}(\mathbb{C}) = GL_3(\mathbb{C}) \times \mathbb{C}^\times$, and the irreducible $\mathcal{V}$ are associated with highest weight representations of the Levi factor $\mathbf{L}(\mathbb{C})$ of $\mathbf{H}(\mathbb{C})$. It is straightforward to check that $\det\mathcal{P}$ and $\mathcal{L}$ are associated with the same character of $\mathbf{L}(\mathbb{C})$, up to a twist by a character of $\mathbf{G}(\mathbb{C})$, which affects the $\mathbf{G}(\mathbb{A}_f)$-action (hence the normalization of Hecke operators), but not the structure of the line bundles themselves. The functoriality of Harris' construction implies that $\det\mathcal{P}$ and $\mathcal{L}$ are isomorphic also algebraically.

We de not know whether $\det\mathcal{P}$ and $\mathcal{L}$ are isomorphic as algebraic line bundles over S. This would be equivalent, by (1.84), to the statement that for every PEL structure $(A, \lambda, \iota, \alpha) \in \mathcal{M}(R)$, for any R_0-algebra R, $\det(H^1_{dR}(A/R)(\Sigma))$ is the trivial line bundle on $\mathrm{Spec}(R)$. To our regret, we have not been able to establish this, although a similar statement in the "Siegel case", namely that for any principally polarized abelian scheme (A, λ) over R, $\det H^1_{dR}(A/R)$ is trivial, follows at once from the Hodge filtration (1.81). Our result, however, suffices to guarantee the following corollary, which is all that we will be using in the sequel.

Corollary 1.18 *For any characteristic p geometric point* $Spec(k) \to Spec(R_0)$, *we have* $\det\mathcal{P} \simeq \mathcal{L}$ *on* $\bar{S}_k$. *A similar statement holds for morphisms* $Spec\,W(k) \to Spec(R_0)$.

Proof Since $\bar{S}$ is a regular scheme, $\det\mathcal{P} \otimes \mathcal{L}^{-1} \simeq \mathcal{O}(D)$ for a Weil divisor D supported on vertical fibers over R_0. Since any connected component Z of $\bar{S}_k$ is irreducible, we can modify D so that D and Z are disjoint, showing that $\det\mathcal{P} \otimes \mathcal{L}^{-1}|_Z$ is trivial. The second claim is proved similarly.

1.7.4 Modular forms

Let R be an R_0-algebra. A *modular form* of weight $k \geq 0$ and level $N \geq 3$ defined over R is an element of the finite R-module

$$M_k(N, R) = H^0\left(\bar{S}_R, \mathcal{L}^k\right). \tag{1.103}$$

We usually omit the subscript R, remembering that $\bar{S}$ is now to be considered over R. The well-known Koecher principle says that $H^0(\bar{S}, \mathcal{L}^k) = H^0(S, \mathcal{L}^k)$. See [2], Section 2.2, for an arithmetic proof valid integrally over any R_0-algebra R. A *cusp form* is an element of the space

$$M_k^0(N,R) = H^0\left(\bar{S}, \mathcal{L}^k \otimes \mathcal{O}(C)^\vee\right). \tag{1.104}$$

As we shall see below (*cf* Corollary 1.23), if $k \geq 3$, there is an isomorphism $\mathcal{L}^k \otimes \mathcal{O}(C)^\vee \simeq \Omega^2_{\bar{S}} \otimes \mathcal{L}^{k-3}$. In particular, cusp forms of weight 3 are "the same" as holomorphic 2-forms on $\bar{S}$.

An alternative definition (à la Katz) of a modular form of weight k and level N defined over R, is as a "rule" f which assigns to every R-scheme T, and every $\underline{A} = (A, \lambda, \iota, \alpha) \in \mathcal{M}(T)$, together with a nowhere vanishing section $\omega \in H^0(T, \omega_{A/T}(\bar{\Sigma}))$, an element $f(\underline{A}, \omega) \in H^0(T, \mathcal{O}_T)$ satisfying

- $f(\underline{A}, \lambda\omega) = \lambda^{-k} f(\underline{A}, \omega)$ for every $\lambda \in H^0(T, \mathcal{O}_T)^\times$
- The "rule" f is compatible with base change T'/T.

Indeed, if f is an element of $M_k(N,R)$, then given such an $\underline{A}$ and ω, the universal property of S produces a unique morphism $\varphi : T \to S$ over R, $\varphi^*\mathcal{A} = A$, and we may let $f(\underline{A}, \omega) = \varphi^* f/\omega^k$. Conversely, given such a rule f we may cover S by Zariski open sets T where $\mathcal{L}$ is trivialized, and then the sections $f(\mathcal{A}_T, \omega_T)\omega_T^k$ (ω_T a trivializing section over T) glue to give $f \in M_k(N,R)$. While viewing f as a "rule" rather than a section is a matter of language, it is sometimes more convenient to use this language.

Let $R \to R'$ be a homomorphism of R_0-algebras. Then Bellaïche proved the following theorem ([2], 1.1.5).

Theorem 1.19 *If $k \geq 3$ (resp. $k \geq 6$), then $M_k^0(N,R)$ (resp. $M_k(N,R)$) is a locally free finite R-module, and the base-change homomorphism*

$$R' \otimes M_k^0(N,R) \simeq M_k^0(N,R') \tag{1.105}$$

is an isomorphism (resp. base change for $M_k(N,R)$).

Bellaïche considers only weights divisible by 3, but his proofs generalize to all k (*cf* remark on the bottom of p. 43 in [2]).

Over $\mathbb{C}$, pulling back to $\mathfrak{X}$ and using the trivialization of $\mathcal{L}$ given by the nowhere vanishing section $2\pi i \cdot d\zeta_3$, a modular form of weight k is a collection $(f_j)_{1\leq j\leq m}$ of holomorphic functions on $\mathfrak{X}$ satisfying

$$f_j(\gamma(z,u)) = j(\gamma; z, u)^k f_j(z,u) \quad \forall \gamma \in \Gamma_j \tag{1.106}$$

(the Koecher principle means that no condition has to be imposed at the cusps).

1.8 The Kodaira–Spencer isomorphism

Let $\pi : A \to S$ be an abelian scheme of relative dimension 3, as in the Picard moduli problem. The Gauss–Manin connection

$$\nabla : H^1_{dR}(A/S) \to H^1_{dR}(A/S) \otimes_{\mathcal{O}_S} \Omega^1_S \tag{1.107}$$

defines the Kodaira–Spencer map

$$\mathrm{KS} \in \mathrm{Hom}_{\mathcal{O}_S}\left(\omega_A \otimes_{\mathcal{O}_S} \omega_{A^t}, \Omega^1_S\right) \tag{1.108}$$

as the composition of the maps

$$\begin{aligned} \omega_A &= H^0(A, \Omega^1_{A/S}) \hookrightarrow H^1_{dR}(A/S) \xrightarrow{\nabla} H^1_{dR}(A/S) \otimes_{\mathcal{O}_S} \Omega^1_S \\ &\twoheadrightarrow R^1\pi_*\mathcal{O}_A \otimes_{\mathcal{O}_S} \Omega^1_S \simeq \omega^\vee_{A^t} \otimes_{\mathcal{O}_S} \Omega^1_S, \end{aligned} \tag{1.109}$$

and finally using $\mathrm{Hom}(L, M^\vee \otimes N) = \mathrm{Hom}(L \otimes M, N)$. Recall that if A is endowed with an $\mathcal{O}_\mathcal{K}$ action via ι, then the induced action of $a \in \mathcal{O}_K$ on A^t is induced from the action on $Pic(A)$, taking a line bundle $\mathcal{M}$ to $\iota(a)^*\mathcal{M}$. As the polarization $\lambda : A \to A^t$ is $\mathcal{O}_S$-linear but satisfies $\lambda \circ \iota(a) = \iota(a^\rho) \circ \lambda$, it follows that the induced $\mathcal{O}_\mathcal{K}$ action on A^t is of type $(1, 2)$ and hence $\omega_{A^t}^\vee$ is of type $(1, 2)$.

Lemma 1.20 *The map KS induces maps*

$$\begin{aligned} \mathrm{KS}(\Sigma) &: \omega_A(\Sigma) \to \omega_{A^t}^\vee(\Sigma) \otimes_{\mathcal{O}_S} \Omega_S^1 \\ \mathrm{KS}(\bar{\Sigma}) &: \omega_A(\bar{\Sigma}) \to \omega_{A^t}^\vee(\bar{\Sigma}) \otimes_{\mathcal{O}_S} \Omega_S^1 \end{aligned} \tag{1.110}$$

hence maps, denoted by the same symbols,

$$\begin{aligned} \mathrm{KS}(\Sigma) &: \omega_A(\Sigma) \otimes_{\mathcal{O}_S} \omega_{A^t}(\Sigma) \to \Omega_S^1 \\ \mathrm{KS}(\bar{\Sigma}) &: \omega_A(\bar{\Sigma}) \otimes_{\mathcal{O}_S} \omega_{A^t}(\bar{\Sigma}) \to \Omega_S^1. \end{aligned} \tag{1.111}$$

The CM-type-reversing isomorphism $\lambda^* : \omega_{A^t} \to \omega_A$ induced by the principal polarization satisfies

$$\mathrm{KS}(\Sigma)(\lambda^* x \otimes y) = \mathrm{KS}(\bar{\Sigma})(\lambda^* y \otimes x) \tag{1.112}$$

for all $x \in \omega_{A^t}(\bar{\Sigma})$ and $y \in \omega_{A^t}(\Sigma)$.

Proof The first claim follows from the fact that the Gauss–Manin connection commutes with the endomorphisms, hence preserves CM types. The second claim is a consequence of the symmetry of the polarization, see [11], Prop. 9.1 on p.81 (in the Siegel modular case).

Observe that $\omega_A(\Sigma) \otimes_{\mathcal{O}_S} \omega_{A^t}(\Sigma)$, as well as $\omega_A(\bar{\Sigma}) \otimes_{\mathcal{O}_S} \omega_{A^t}(\bar{\Sigma})$, are vector bundles of rank 2.

Lemma 1.21 *If S is the Picard modular surface and $A = \mathcal{A}$ is the universal abelian variety, then*

$$\mathrm{KS}(\Sigma) : \omega_\mathcal{A}(\Sigma) \otimes_{\mathcal{O}_S} \omega_{\mathcal{A}^t}(\Sigma) \to \Omega_S^1 \tag{1.113}$$

is an isomorphism, and so is $\mathrm{KS}(\bar{\Sigma})$.

Proof This is well known and follows from deformation theory. For a self-contained proof, see [2], Prop. II.2.1.5.

Proposition 1.22 *The Kodaira–Spencer map induces a canonical isomorphism of vector bundles over S*

$$\mathcal{P} \otimes \mathcal{L} \simeq \Omega_S^1. \tag{1.114}$$

Proof We need only use λ^* to identify $\omega_{\mathcal{A}^t}(\Sigma)$ with $\omega_\mathcal{A}(\bar{\Sigma})$.

We refer to Corollary 1.29 for an extension of this result to $\bar{S}$.

Corollary 1.23 *There is an isomorphism of line bundles $\mathcal{L}^3 \simeq \Omega_S^2$.*

Proof Take determinants and use $\det\mathcal{P} \simeq \mathcal{L}$. We emphasize that while $\mathrm{KS}(\Sigma)$ is canonical, the identification of $\det\mathcal{P}$ with $\mathcal{L}$ depends on a choice, which we shall fix later on once and for all.

The last corollary should be compared to the case of the open modular curve $Y(N)$, where the *square* of the Hodge bundle $\omega_{\mathcal{E}}$ of the universal elliptic curve becomes isomorphic to $\Omega^1_{Y(N)}$. Over $\mathbb{C}$, as the isomorphism between $\mathcal{L}^3$ and Ω^2_S takes $\mathrm{d}\zeta_3^{\otimes 3}$ to a constant multiple of $\mathrm{d}z \wedge \mathrm{d}u$ (see Corollary 1.31), the differential form corresponding to a modular form $(f_j)_{1\le j\le m}$ of weight 3, is (up to a constant) $(f_j(z,u)\mathrm{d}z \wedge \mathrm{d}u)_{1\le j\le m}$.

1.9 Extensions to the boundary of *S*

1.9.1 The vector bundles $\mathcal{P}$ and $\mathcal{L}$ over C

Let $E \subset C_{R_N}$ be a connected component of the cuspidal divisor (over the integral closure R_N of R_0 in the ray class field $\mathcal{K}_N$). As we have seen, E is an elliptic curve with CM by $\mathcal{O}_{\mathcal{K}}$. If the cusp at which E sits is of type $(\mathfrak{a}, B)$ ($\mathfrak{a}$ an ideal of $\mathcal{O}_{\mathcal{K}}$, B an elliptic curve with CM by $\mathcal{O}_{\mathcal{K}}$ defined over R_N), then E maps via an isogeny to $\delta_{\mathcal{K}}\mathfrak{a}\otimes_{\mathcal{O}_{\mathcal{K}}} B^t = \mathrm{Ext}^1_{\mathcal{O}_{\mathcal{K}}}(B, \mathfrak{a}\otimes\mathbb{G}_m)$. In particular, E and B are isogenous over $\mathcal{K}_N$.

Consider $\mathcal{G}$, the universal semi-abelian $\mathcal{O}_{\mathcal{K}}$-threefold of type $(\mathfrak{a}, B)$, over $\delta_{\mathcal{K}}\mathfrak{a}\otimes_{\mathcal{O}_{\mathcal{K}}} B^t$. The semi-abelian scheme $\mathcal{A}$ over E is the pull-back of this $\mathcal{G}$. Clearly, $\omega_{\mathcal{A}/E} = \mathcal{P}\oplus\mathcal{L}$ and $\mathcal{P} = \omega_{\mathcal{A}/E}(\Sigma)$ admits over E a canonical rank 1 sub-bundle $\mathcal{P}_0 = \omega_B$. Since the toric part and the abelian part of $\mathcal{G}$ are constant, $\mathcal{L}$, $\mathcal{P}_0$ and $\mathcal{P}_\mu = \mathcal{P}/\mathcal{P}_0$ are all trivial line bundles when restricted to E. It can be shown that $\mathcal{P}$ itself is not trivial over E.

1.9.2 More identities over $\bar{S}$

We have seen that $\Omega^2_S \simeq \mathcal{L}^3$. For the following proposition, compare [2], Lemme II.2.1.7.

Proposition 1.24 *Working over $\mathcal{K}_N$, let E_j $(1\le j\le h)$ be the connected components of C. Then*

$$\Omega^2_{\bar{S}} \simeq \mathcal{L}^3 \otimes \bigotimes_{j=1}^{h} \mathcal{O}(E_j)^\vee. \tag{1.115}$$

Proof By [15] II.6.5, $\Omega^2_{\bar{S}} \simeq \mathcal{L}^3\otimes\bigotimes_{j=1}^h \mathcal{O}(E_j)^{n_j}$ for some integers n_j and we want to show that $n_j = -1$ for all j. By the adjunction formula on the smooth surface $\bar{S}$, if we denote by $K_{\bar{S}}$ a canonical divisor, $\mathcal{O}(K_{\bar{S}}) = \Omega^2_{\bar{S}}$, then

$$0 = 2g_{E_j} - 2 = E_j\cdot(E_j + K_{\bar{S}}). \tag{1.116}$$

We conclude that

$$\deg\left(\Omega^2_{\bar{S}}|_{E_j}\right) = E_j\cdot K_{\bar{S}} = -E_j\cdot E_j > 0. \tag{1.117}$$

Here $E_j\cdot E_j < 0$ because E_j can be contracted to a point (Grauert's theorem). As $\mathcal{L}|_{E_j}$ and $\mathcal{O}(E_i)|_{E_j}$ $(i\ne j)$ are trivial, we get

$$-E_j\cdot E_j = n_j E_j\cdot E_j, \tag{1.118}$$

hence $n_j = -1$ as desired.

1.10 Fourier–Jacobi expansions

1.10.1 The infinitesimal retraction

We follow the arithmetic theory of Fourier–Jacobi expansions as developed in [2]. Let $\widehat{S}$ be the formal completion of $\bar{S}$ along the cuspidal divisor $C = \bar{S}\backslash S$. We work over R_0 and denote by $C^{(n)}$ the n-th infinitesimal neighborhood of C in $\bar{S}$. The closed immersion $i : C \hookrightarrow \widehat{S}$ admits a canonical left inverse $r : \widehat{S} \to C$, a *retraction* satisfying $r \circ i = Id_C$. This is not automatic, but rather a consequence of the rigidity of tori, as explained in [2], Proposition II.2.4.2. As a corollary, the universal semi-abelian scheme $\mathcal{A}_{/C^{(n)}}$ is the pull-back of $\mathcal{A}_{/C}$ via r. The same therefore holds for $\mathcal{P}$ and $\mathcal{L}$, namely there are natural isomorphisms $r^*(\mathcal{P}|_C) \simeq \mathcal{P}|_{C^{(n)}}$ and $r^*(\mathcal{L}|_C) \simeq \mathcal{L}|_{C^{(n)}}$. As a consequence, the filtration

$$0 \to \mathcal{P}_0 \to \mathcal{P} \to \mathcal{P}_\mu \to 0 \tag{1.119}$$

extends canonically to $C^{(n)}$. Since $\mathcal{L}, \mathcal{P}_0$ and $\mathcal{P}_\mu$ are trivial on C, they are trivial over $C^{(n)}$ as well.

1.10.2 Arithmetic Fourier–Jacobi expansions

We fix an arbitrary Noetherian R_0-algebra R and consider all our schemes over R, without a change in notation. As usual, we let $\mathcal{O}_{\widehat{S}} = \lim_{\leftarrow} \mathcal{O}_{C^{(n)}}$ (a sheaf in the Zariski topology on C). Via r^*, this is a sheaf of $\mathcal{O}_C$-modules. Choose a global nowhere vanishing section $s \in H^0(C, \mathcal{L})$ trivializing $\mathcal{L}$. Such a section is unique up to a unit of R on each connected component of C. This s determines an isomorphism

$$\mathcal{L}^k|_{\widehat{S}} \simeq \mathcal{O}_{\widehat{S}}, \quad f \mapsto f/(r^*s)^k \tag{1.120}$$

for each k, hence a ring homomorphism

$$FJ : \oplus_{k=0}^{\infty} M_k(N, R) \to H^0(C, \mathcal{O}_{\widehat{S}}). \tag{1.121}$$

We call $FJ(f)$ the (arithmetic) Fourier–Jacobi expansion of f. It depends on s in an obvious way.

To understand the structure of $H^0(C, \mathcal{O}_{\widehat{S}})$ let $\mathcal{I} \subset \mathcal{O}_{\bar{S}}$ be the sheaf of ideals defining C, so that $C^{(n)}$ is defined by $\mathcal{I}^n$. The conormal sheaf $\mathcal{N} = \mathcal{I}/\mathcal{I}^2$ is the restriction $i^*\mathcal{O}_{\bar{S}}(-C)$ of $\mathcal{I} = \mathcal{O}_{\bar{S}}(-C)$ to C. It is an ample invertible sheaf on C, since (over R_N) its degree on each component E_j is $-E_j^2 > 0$.

Now r^* supplies, for every $n \geq 2$, a canonical splitting of

$$0 \to \mathcal{I}/\mathcal{I}^n \to \mathcal{O}_{\bar{S}}/\mathcal{I}^n \leftrightarrows \mathcal{O}_{\bar{S}}/\mathcal{I} \to 0. \tag{1.122}$$

Inductively, we get a direct sum decomposition

$$\mathcal{O}_{\bar{S}}/\mathcal{I}^n \simeq \bigoplus_{m=0}^{n-1} \mathcal{I}^m/\mathcal{I}^{m+1} \tag{1.123}$$

as $\mathcal{O}_C$-modules, hence, since $\mathcal{I}^m/\mathcal{I}^{m+1} \simeq \mathcal{N}^m$, an isomorphism

$$H^0(C, \mathcal{O}_{C^{(n)}}) \simeq \bigoplus_{m=0}^{n-1} H^0(C, \mathcal{N}^m), \quad f \mapsto \sum_{m=0}^{n-1} c_m(f). \tag{1.124}$$

This isomorphism respects the multiplicative structure, so is a ring isomorphism. Going to the projective limit, and noting that the $c_m(f)$ are independent of n, we get

$$FJ(f) = \sum_{m=0}^{\infty} c_m(f) \in \prod_{m=0}^{\infty} H^0(C, \mathcal{N}^m). \tag{1.125}$$

1.10.3 Fourier–Jacobi expansions over $\mathbb{C}$

Working over $\mathbb{C}$, we shall now relate the infinitesimal retraction r to the geodesic retraction, and the powers of the conormal bundle $\mathcal{N}$ to theta functions. Recall the analytic compactification of X_Γ described in Proposition 1.7. Let E be the connected component of $\bar{X}_\Gamma \backslash X_\Gamma$ corresponding to the standard cusp c_∞. As before, we denote by $E^{(n)}$ its nth infinitesimal neighborhood. The line bundle $\mathcal{T}|_E$ is just the analytic normal bundle to E, and hence we have an isomorphism

$$\mathcal{N}_{an} \simeq \mathcal{T}^\vee \tag{1.126}$$

between the analytification of $\mathcal{N} = \mathcal{I}/\mathcal{I}^2$ and the dual of $\mathcal{T}$.

Lemma 1.25 *The infinitesimal retraction $r : E^{(n)} \to E$ coincides with the map induced by the geodesic retraction* (1.15)*.*

Proof The meaning of the lemma is this. The infinitesimal retraction induces a map of ringed spaces

$$r_{an} : E_{an}^{(n)} \to E_{an} \tag{1.127}$$

where E_{an} is the analytic space associated with E with its sheaf of analytic functions $\mathcal{O}_E^{hol}$, and $E_{an}^{(n)}$ is the same topological space with the sheaf $\mathcal{O}_{\bar{S}}^{hol}/\mathcal{I}_{an}^n$. The geodesic retraction (sending (z, u) to $u \mod \Lambda$) is an analytic map $r_{geo} : E_{an}(\varepsilon) \to E_{an}$, where $E_{an}(\varepsilon)$ is our notation for some tubular neighborhood of E_{an} in $\bar{S}_{an}$. On the other hand, there is a canonical map *can* of ringed spaces from $E_{an}^{(n)}$ to $E_{an}(\varepsilon)$. We claim that these three maps satisfy $r_{geo} \circ can = r_{an}$.

To prove the lemma, note that the infinitesimal retraction $r : E^{(n)} \to E$ is uniquely characterized by the fact that the $\mathcal{O}_\mathcal{K}$-semi-abelian variety $\mathcal{A}_x = x^*\mathcal{A}$ at any point $x : \mathrm{Spec}(R) \to E^{(n)}$ is equal to $\mathcal{A}_{r\circ x}$ (an equality respecting the PEL structures). See [2], II.2.4.2. The computations of Sect. 1.6 show that the same is true for the infinitesimal retraction obtained from the geodesic retraction. We conclude that the two retractions agree on the level of "truncated Taylor expansions".

Consider now a modular form of weight k and level N over $\mathbb{C}$, $f \in M_k(N, \mathbb{C})$. Using the trivialization of $\mathcal{L}_{an}$ over the symmetric space $\mathfrak{X}$ given by $2\pi i \cdot \mathrm{d}\zeta_3$ as discussed in Sect. 1.7.2, we identify f with a collection of *functions* f_j on $\mathfrak{X}$, transforming under Γ_j according to the automorphy factor $j(\gamma; z, u)^k$. As usual, we look at $\Gamma = \Gamma_1$ only, and at the expansion of $f = f_1$ at the standard cusp c_∞, the other cusps being in principle similar. On the arithmetic FJ expansion side, this means that we concentrate on one connected component E of C, which lies on the connected component of $S_\mathbb{C}$ corresponding to $g_1 = 1$. It also means that as the section s used to trivialize $\mathcal{L}$ along E, we must use a section that, analytically, coincides with $2\pi i \cdot \mathrm{d}\zeta_3$.

Pulling back the sheaf $\mathcal{N}_{an}$ from $E = \mathbb{C}/\Lambda$ to $\mathbb{C}$, it is clear that $q = q(z) = e^{2\pi i z/M}$ maps, at each $u \in \mathbb{C}$, to a generator of $\mathcal{T}^\vee = \mathcal{N}_{an} = \mathcal{I}_{an}/\mathcal{I}_{an}^2$, and we denote by q^m the corresponding generator of $\mathcal{N}_{an}^m = \mathcal{I}_{an}^m/\mathcal{I}_{an}^{m+1}$. If

$$f(z, u) = \sum_{m=0}^{\infty} \theta_m(u) e^{2\pi i m z/M} = \sum_{m=0}^{\infty} \theta_m(u) q^m \tag{1.128}$$

is the complex analytic Fourier expansion of f at a neighborhood of c_∞, then $c_m(z, u) = \theta_m(u)q^m \in H^0(E, \mathcal{N}_{an}^m)$ is just the restriction of the section denoted above by $c_m(f)$ to E. The functions θ_m are classical elliptic theta functions (for the lattice Λ).

1.11 The Gauss–Manin connection in a neighborhood of a cusp

1.11.1 A computation of ∇ in the complex model

We shall now compute the Gauss–Manin connection in the complex model near the standard cusp c_∞. Recall that we use the coordinates $(z, u, \zeta_1, \zeta_2, \zeta_3)$ as in Sect. 1.2.4. Here $\mathrm{d}\zeta_1$ and $\mathrm{d}\zeta_2$ form a basis for $\mathcal{P}$ and $\mathrm{d}\zeta_3$ for $\mathcal{L}$. The same coordinates served to define also the semi-abelian variety $\mathcal{G}_u$ (denoted also $\mathcal{A}_u$) over the cuspidal component E at c_∞, *cf* Sect. 1.6. As explained there (1.69), the projection to the abelian part is given by the coordinate ζ_1 (modulo $\mathcal{O}_\mathcal{K}$), so $\mathrm{d}\zeta_1$ is a basis for the sub-line-bundle of $\omega_{\mathcal{A}/E}$ coming from the abelian part, which was denoted $\mathcal{P}_0$. In Sect. 1.10.1, it is explained how to extend the filtration $\mathcal{P}_0 \subset \mathcal{P}$ canonically to the formal neighborhood $\widehat{S}$ of E using the retraction r, by pulling back from the boundary. It was also noted that complex analytically, the retraction r is the germ of the geodesic retraction introduced earlier. From the analytic description of the degeneration of $\mathcal{A}_{(z,u)}$ along a geodesic, it becomes clear that $\mathcal{P}_0 = r^*(\mathcal{P}_0|_E)$ is just the line bundle $\mathcal{O}_{\widehat{S}} \cdot \mathrm{d}\zeta_1 \subset \omega_{\mathcal{A}/\widehat{S}}$. It follows that $\mathcal{P}_\mu = \mathcal{O}_{\widehat{S}} \cdot \mathrm{d}\zeta_2 \mod \mathcal{P}_0$.

We shall now pull back these vector bundles to $\mathfrak{X}$ and compute the Gauss Manin connection ∇ complex analytically on $\omega_{\mathcal{A}/\mathfrak{X}}$. We write $\mathcal{P}_0 = \mathcal{O}_\mathfrak{X} \cdot \mathrm{d}\zeta_1$ for $\mathcal{P}_{0,an}$ etc. dropping the decoration *an.* Recalling that $\mathcal{O}_\mathcal{K} = \mathbb{Z} \oplus \mathbb{Z}\omega_\mathcal{K}$, we let

$$\alpha_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad \alpha_2 = \begin{pmatrix} 1 \\ 0 \\ u \end{pmatrix}, \quad \alpha_3 = \begin{pmatrix} u \\ -z/\delta \\ z/\delta \end{pmatrix} \tag{1.129}$$

and

$$\alpha_1' = \iota'(\omega_\mathcal{K})\alpha_1 = \begin{pmatrix} 0 \\ \omega_\mathcal{K} \\ \bar{\omega}_\mathcal{K} \end{pmatrix}, \quad \alpha_2' = \iota'(\omega_\mathcal{K})\alpha_2 = \begin{pmatrix} \omega_\mathcal{K} \\ 0 \\ \bar{\omega}_\mathcal{K} u \end{pmatrix},$$
$$\alpha_3' = \iota'(\omega_\mathcal{K})\alpha_3 = \begin{pmatrix} \omega_\mathcal{K} u \\ -\omega_\mathcal{K} z/\delta \\ \bar{\omega}_\mathcal{K} z/\delta \end{pmatrix}. \tag{1.130}$$

These 6 vectors span $L'_{(z,u)}$ over $\mathbb{Z}$. Let $\beta_1, \ldots, \beta_3'$ be the dual basis to $\{\alpha_1, \ldots, \alpha_3'\}$ in $H^1_{dR}(\mathcal{A}/\mathcal{O}_\mathfrak{X})$, i.e. $\int_{\alpha_1} \beta_1 = 1$ etc. As the periods of the β_i's along the integral homology are constant, the β-basis is horizontal for the Gauss–Manin connection. The first coordinate of the α_i and α_i' gives us

$$\mathrm{d}\zeta_1 = 0 \cdot \beta_1 + 1 \cdot \beta_2 + u \cdot \beta_3 + 0 \cdot \beta_1' + \omega_\mathcal{K} \cdot \beta_2' + \omega_\mathcal{K} u \cdot \beta_3', \tag{1.131}$$

and we find that

$$\nabla(\mathrm{d}\zeta_1) = \left(\beta_3 + \omega_\mathcal{K}\beta_3'\right) \otimes \mathrm{d}u. \tag{1.132}$$

Similarly, we find

$$\nabla(\mathrm{d}\zeta_2) = -\delta^{-1}\left(\beta_3 + \omega_\mathcal{K}\beta_3'\right) \otimes \mathrm{d}z \tag{1.133}$$
$$\nabla(\mathrm{d}\zeta_3) = \left(\beta_2 + \bar{\omega}_\mathcal{K}\beta_2'\right) \otimes \mathrm{d}u + \delta^{-1}\left(\beta_3 + \bar{\omega}_\mathcal{K}\beta_3'\right) \otimes \mathrm{d}z.$$

1.11.2 A computation of KS in the complex model

We go on to compute the Kodaira–Spencer map on $\mathcal{P}$, i.e. the map denoted $\mathrm{KS}(\Sigma)$. For that, we have to take $\nabla(\mathrm{d}\zeta_1)$ and $\nabla(\mathrm{d}\zeta_2)$ and project them to $R^1\pi_*\mathcal{O}_\mathcal{A}(\Sigma) \otimes \Omega^1_\mathfrak{X}$. We then pair the result, using the polarization form $\langle , \rangle_\lambda$ on $H^1_{dR}(\mathcal{A})$ (reflecting the isomorphism

$$R^1\pi_*\mathcal{O}_{\mathcal{A}}(\Sigma) = Lie(\mathcal{A}^t)(\Sigma) = \omega^\vee_{\mathcal{A}^t}(\Sigma) \simeq \mathcal{L}^\vee(\rho) \tag{1.134}$$

coming from λ), with $\mathrm{d}\zeta_3$.

To perform the computation, we need two lemmas.

Lemma 1.26 *The Riemann form on L'_x, associated with the polarization λ, is given in the basis $\alpha_1, \alpha_2, \alpha_3, \alpha'_1, \alpha'_2, \alpha'_3$ by the matrix*

$$J = \begin{pmatrix} & & & & & 1 \\ & & & & -1 & \\ & & & 1 & & \\ & & -1 & & & \\ & 1 & & & & \\ -1 & & & & & \end{pmatrix}. \tag{1.135}$$

Proof This is an easy computation using the transition map T between L and L'_x and the fact that on L the Riemann form is the alternating form $\langle , \rangle = \mathrm{Im}_\delta(,)$.

For the formulation of the next lemma recall that if A is a complex abelian variety, a polarization $\lambda : A \to A^t$ induces an alternating form $\langle , \rangle_\lambda$ on $H^1_{dR}(A)$ as well as a Riemann form on the integral homology $H_1(A, \mathbb{Z})$. We compare the two.

Lemma 1.27 *Let (A, λ) be a principally polarized complex abelian variety. If $\alpha_1, \ldots, \alpha_{2g}$ is a symplectic basis for $H_1(A, \mathbb{Z})$ in which the associated Riemann form is given by a matrix J, and $\beta_1, \ldots, \beta_{2g}$ is the dual basis of $H^1_{dR}(A)$, then the matrix of the bilinear form $\langle , \rangle_\lambda$ on $H^1_{dR}(A)$ in the basis $\beta_1, \ldots, \beta_{2g}$ is $(2\pi i)^{-1}J$.*

Proof These are essentially Riemann's bilinear relations. For example, if A is the Jacobian of a curve $\mathcal{C}$ and the basis $\alpha_1, \ldots, \alpha_{2g}$ has the standard intersection matrix

$$J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \tag{1.136}$$

then the lemma follows from the well-known formula for the cup product (ξ, η being differentials of the second kind on $\mathcal{C}$)

$$\xi \cup \eta = \frac{1}{2\pi i} \sum_{i=1}^{g} \left(\int_{\alpha_i} \xi \int_{\alpha_{i+g}} \eta - \int_{\alpha_i} \eta \int_{\alpha_{i+g}} \xi \right). \tag{1.137}$$

Using the two lemmas, we get

$$\begin{aligned} \mathrm{KS}(\mathrm{d}\zeta_1 \otimes \mathrm{d}\zeta_3) &= \langle \beta_3 + \omega_{\mathcal{K}}\beta'_3, \mathrm{d}\zeta_3 \rangle_\lambda \cdot \mathrm{d}u \\ &= \begin{matrix} \langle \beta_3 + \omega_{\mathcal{K}}\beta'_3,\ \beta_1 + u\beta_2 + z\delta^{-1}\beta_3 + \\ \bar{\omega}_{\mathcal{K}}\beta'_1 + \bar{\omega}_{\mathcal{K}}u\beta'_2 + \bar{\omega}_{\mathcal{K}}z\delta^{-1}\beta'_3 \rangle_\lambda \cdot \mathrm{d}u \end{matrix} \\ &= -\delta(2\pi i)^{-1}\mathrm{d}u. \end{aligned} \tag{1.138}$$

Similarly,

$$\begin{aligned} \mathrm{KS}(\mathrm{d}\zeta_2 \otimes \mathrm{d}\zeta_3) &= \langle -\delta^{-1}(\beta_3 + \omega_{\mathcal{K}}\beta'_3), \mathrm{d}\zeta_3 \rangle_\lambda \cdot \mathrm{d}z \\ &= (2\pi i)^{-1}\mathrm{d}z. \end{aligned} \tag{1.139}$$

We summarize.

Proposition 1.28 *Let $z, u, \zeta_1, \zeta_2, \zeta_3$ be the standard coordinates in a neighborhood of the cusp c_∞. Then, complex analytically, the Kodaira–Spencer isomorphism*

$$\mathrm{KS}(\Sigma) : \mathcal{P} \otimes \mathcal{L} \simeq \Omega^1_{\mathfrak{X}} \tag{1.140}$$

is given by the formulae

$$\mathrm{KS}(\mathrm{d}\zeta_1 \otimes \mathrm{d}\zeta_3) = -\delta(2\pi i)^{-1}\mathrm{d}u, \quad \mathrm{KS}(\mathrm{d}\zeta_2 \otimes \mathrm{d}\zeta_3) = (2\pi i)^{-1}\mathrm{d}z. \tag{1.141}$$

Corollary 1.29 *The Kodaira–Spencer isomorphism $\mathcal{P} \otimes \mathcal{L} \simeq \Omega^1_S$ extends meromorphically over $\bar{S}$. Moreover, in a formal neighborhood $\widehat{S}$ of C, its restriction to the line sub-bundle $\mathcal{P}_0 \otimes \mathcal{L}$ is holomorphic, and on any direct complement of $\mathcal{P}_0 \otimes \mathcal{L}$ in $\mathcal{P} \otimes \mathcal{L}$, it has a simple pole along C.*

Proof As we have seen, $\mathrm{d}\zeta_1 \otimes \mathrm{d}\zeta_3$ and $\mathrm{d}\zeta_2 \otimes \mathrm{d}\zeta_3$ define a basis of $\mathcal{P} \otimes \mathcal{L}$ at the boundary, with $\mathrm{d}\zeta_1 \otimes \mathrm{d}\zeta_3$ spanning the line sub-bundle $\mathcal{P}_0 \otimes \mathcal{L}$. On the other hand $\mathrm{d}u$ is holomorphic there, while $\mathrm{d}z$ has a simple pole along the boundary.

Corollary 1.30 *The induced map*

$$\psi : \Omega^1_{\mathfrak{X}} \to \mathcal{P}_\mu \otimes \mathcal{L} \tag{1.142}$$

($\mathcal{P}_\mu = \mathcal{P}/\mathcal{P}_0$) obtained by inverting the isomorphism $\mathrm{KS}(\Sigma)$ *and dividing $\mathcal{P}$ by $\mathcal{P}_0$ kills* $\mathrm{d}u$ *and maps* $\mathrm{d}z$ *to* $2\pi i \cdot \mathrm{d}\zeta_2 \otimes \mathrm{d}\zeta_3$.

Proof As we have seen, $\mathrm{d}\zeta_1$ is a basis for $\mathcal{P}_0$.

Corollary 1.31 *The isomorphism $\mathcal{L}^3 \simeq \Omega^2_S$ maps $\mathrm{d}\zeta_3^{\otimes 3}$ to a constant multiple of $\mathrm{d}z \wedge \mathrm{d}u$.*

Proof The isomorphism $\det \mathcal{P} \simeq \mathcal{L}$ carries $\mathrm{d}\zeta_1 \wedge \mathrm{d}\zeta_2$ to a constant multiple of $\mathrm{d}\zeta_3$, so the corollary follows from (1.141).

1.11.3 Transferring the results to the algebraic category

The computations leading to (1.141) of course descend (still in the analytic category) to $S_{\mathbb{C}}$, because they are local in nature. They then hold *a fortiori* in the formal completion $\widehat{S}_{\mathbb{C}}$ along the cuspidal component E. Recall that the sections $\mathrm{d}\zeta_1$, $\mathrm{d}\zeta_3$ and $\mathrm{d}\zeta_2 \mod \langle \mathrm{d}\zeta_1 \rangle$ (respectively, $\mathrm{d}u$ and $\mathrm{d}z \mod \langle \mathrm{d}u \rangle$) are well defined in $\widehat{S}_{\mathbb{C}}$, because as global sections defined over $\mathfrak{X}$ they are invariant under Γ_{cusp} (see Lemma 1.8). But the Gauss–Manin and Kodaira–Spencer maps are defined algebraically on S, and both $\Omega^1_{\widehat{S}}$ and $\omega_{\mathcal{A}/\widehat{S}}$ are flat over R_0, so from the validity of the formulae over $\mathbb{C}$ we deduce their validity in $\widehat{S}$ over R_0, provided we identify the differential forms figuring in them (suitably normalized) with elements of $\Omega^1_{\widehat{S}}$ and $\omega_{\mathcal{A}/\widehat{S}}$ defined over R_0. In particular, they hold in the characteristic p fiber as well.

From the relation

$$\frac{\mathrm{d}q}{q} = \frac{2\pi i}{M}\mathrm{d}z, \tag{1.143}$$

we deduce that the map ψ has a simple zero along the cuspidal divisor.

Finally, although we have done all the computations at one specific cusp, it is clear that similar computations hold at any other cusp.

1.12 Fields of rationality

1.12.1 Rationality of local sections of $\mathcal{P}$ and $\mathcal{L}$

We have compared the arithmetic surface S with the complex analytic surfaces $\Gamma_j\backslash\mathfrak{X}$ ($1 \le j \le m$), and the compactifications of these two models. We have also compared the universal semi-abelian scheme $\mathcal{A}$ and the automorphic vector bundles $\mathcal{P}$ and $\mathcal{L}$ in both models. In this section, we want to compare the local parameters obtained from the two presentations, and settle the question of rationality. For simplicity, we shall work rationally and not integrally, which is all we need. In order to work integrally, one would have to study degeneration and periods of abelian varieties integrally, which is more delicate, see [28], Ch.I, Sections 3,4.

We shall need to look at local parameters at the cusps, and as the cusps are defined only over $\mathcal{K}_N$, we shall work with $S_{\mathcal{K}_N}$ instead of $S_{\mathcal{K}}$. With a little more care, working with Galois orbits of cusps, we could probably prove rationality over $\mathcal{K}$, but for our purpose $\mathcal{K}_N$ is good enough.

If ξ and η belong to a $\mathcal{K}_N$-module, we write $\xi \sim \eta$ to mean that $\eta = c\xi$ for some $c \in \mathcal{K}_N^\times$.

We begin with the vector bundles $\mathcal{P}$ and $\mathcal{L}$. Over $\mathbb{C}$, they yield analytic vector bundles $\mathcal{P}_{an}$ and $\mathcal{L}_{an}$ on each X_{Γ_j} ($1 \le j \le m$). Assume for the rest of this section that $j = 1$ and write $\Gamma = \Gamma_1$. Similar results will hold for every j. The vector bundles $\mathcal{P}$ and $\mathcal{L}$ are trivialized over the unit ball $\mathfrak{X}$ by means of the nowhere vanishing sections $\mathrm{d}\zeta_3 \in H^0(\mathfrak{X}, \mathcal{L}_{an})$ and $\mathrm{d}\zeta_1, \mathrm{d}\zeta_2 \in H^0(\mathfrak{X}, \mathcal{P}_{an})$. These sections do not descend to X_Γ, but

$$\sigma_{an} = (\mathrm{d}\zeta_1 \wedge \mathrm{d}\zeta_2) \otimes \mathrm{d}\zeta_3^{-1} \in H^0\left(X_\Gamma, \det\mathcal{P} \otimes \mathcal{L}^{-1}\right) \tag{1.144}$$

does, as the factors of automorphy of $\mathrm{d}\zeta_1 \wedge \mathrm{d}\zeta_2$ and $\mathrm{d}\zeta_3$ are the same (*cf* Sect. 1.7.2). Furthermore, this factor of automorphy (i.e. $j(\gamma; z, u)$) is trivial on Γ_{cusp}, the stabilizer of c_∞ in Γ, so $\mathrm{d}\zeta_1 \wedge \mathrm{d}\zeta_2$ and $\mathrm{d}\zeta_3$ define sections of $\det\mathcal{P}$ and $\mathcal{L}$ on $\widehat{S}_{\mathbb{C}}$, the formal completion of $\bar{S}_{\mathbb{C}}$ along the cuspidal divisor $E_c = p^{-1}(c_\infty) \subset \bar{S}_{\mathbb{C}}$. The same also holds for $\mathrm{d}\zeta_1$ and $\mathrm{d}\zeta_2$ mod $\langle \mathrm{d}\zeta_1\rangle$ individually (Lemma 1.8). Along E_c, $\mathcal{P}$ has a canonical filtration

$$0 \to \mathcal{P}_0 \to \mathcal{P} \to \mathcal{P}_\mu \to 0 \tag{1.145}$$

and $\mathrm{d}\zeta_1$ is a generator of $\mathcal{P}_0$. (Compare (1.63) and (1.71) and note that the projection to $\mathbb{C}/\mathcal{O}_{\mathcal{K}} = B(\mathbb{C})$ is via the coordinate ζ_1, so $\mathrm{d}\zeta_1$ is a generator of $\mathcal{P}_0|_{E_c} = \omega_B$.) As we have shown in Sect. 1.10.1, this filtration extends to the formal neighborhood $\widehat{S}_{\mathbb{C}}$ of E_c. The vector bundles $\mathcal{P}$ and $\mathcal{L}$, as well as the filtration on $\mathcal{P}$, are defined over $\mathcal{K}_N$. It makes sense therefore to ask whether certain sections are $\mathcal{K}_N$-rational. Recall that the cusp c_∞ is of type $(\mathcal{O}_{\mathcal{K}}, \mathcal{O}_{\mathcal{K}})$.

Proposition 1.32 (i) *$2\pi i \cdot \mathrm{d}\zeta_3 \in H^0(\widehat{S}_{\mathcal{K}_N}, \mathcal{L})$. In other words, this section is $\mathcal{K}_N$-rational.*

(ii) *Similarly $2\pi i \cdot \mathrm{d}\zeta_2$ projects (modulo $\mathcal{P}_0$) to a $\mathcal{K}_N$-rational section of $\mathcal{P}_\mu$.*

(iii) *Let B be the elliptic curve over $\mathcal{K}_N$ associated with the cusp c_∞ as in Sect. 1.5.1. Let $\Omega_B \in \mathbb{C}^\times$ be a period of a basis ω of $\omega_B = H^0(B, \Omega^1_{B/\mathcal{K}_N})$ (i.e. the lattice of periods of ω is $\Omega_B \cdot \mathcal{O}_{\mathcal{K}}$). This Ω_B is well defined up to an element of $\mathcal{K}_N^\times$. Then $\Omega_B \cdot d\zeta_1 \in H^0(\widehat{S}_{\mathcal{K}_N}, \mathcal{P}_0)$ is $\mathcal{K}_N$-rational.*

Proof Let E be the component of $C_{\mathcal{K}_N}$ which over $\mathbb{C}$ becomes E_c. Let $\mathcal{G}$ be the universal semi-abelian scheme over E. Then $\mathcal{G}$ is a semi-abelian scheme which is an extension of $B \times_{\mathcal{K}_N} E$ by the torus $(\mathcal{O}_{\mathcal{K}} \otimes \mathbb{G}_{m,\mathcal{K}_N}) \times_{\mathcal{K}_N} E$. At any point $u \in E(\mathbb{C})$, we have the analytic model $\mathcal{G}_u$ (1.69) for the fiber of $\mathcal{G}$ at u, but the abelian part and the toric part are constant.

Over E, the line bundle $\mathcal{P}_0$ is (by definition) $\omega_{B\times E/E}$. As the lattice of periods of a suitable $\mathcal{K}_N$-rational differential is $\Omega_B \cdot \mathcal{O}_\mathcal{K}$, while the lattice of periods of $\mathrm{d}\zeta_1$ is $\mathcal{O}_\mathcal{K}$, part (iii) follows. For parts (i) and (ii), observe that the toric part of $\mathcal{G}$ is in fact defined over $\mathcal{K}$ and that $e^*_{\mathcal{O}_\mathcal{K}}$ maps the cotangent space of $\mathcal{O}_\mathcal{K} \otimes \mathbb{G}_{m,\mathcal{K}}$ isomorphically to the $\mathcal{K}$-span of $2\pi i d\zeta_2$ and $2\pi i d\zeta_3$.

Corollary 1.33 *$\Omega_B \cdot \sigma_{an}$ is a nowhere vanishing global section of* $\det \mathcal{P} \otimes \mathcal{L}^{-1}$ *over S_Γ, rational over $\mathcal{K}_N$.*

Proof Recall that we denote by S_Γ the connected component of $S_{\mathcal{K}_N}$ whose associated analytic space is the complex manifold X_Γ. We have seen that as an analytic section $\Omega_B \cdot \sigma_{an}$ descends to X_Γ and extends to the smooth compactification $\bar{X}_\Gamma$. By GAGA, it is algebraic. Since $\bar{X}_\Gamma$ is connected, to check its field of definition, it is enough to consider it at one of the cusps. By the Proposition, its restriction to the formal neighborhood of E_c ($c = c_\infty$) is defined over $\mathcal{K}_N$.

The complex periods Ω_B (and their powers) appear as the transcendental parts of special values of L-functions associated with Grossencharacters of $\mathcal{K}$. They are therefore instrumental in the construction of p-adic L-functions on $\mathcal{K}$. We expect them to appear in the p-adic interpolation of holomorphic Eisenstein series on the group $\mathbf{G}$, much as powers of $2\pi i$ (values of $\zeta(2k)$) appear in the p-adic interpolation of Eisenstein series on $GL_2(\mathbb{Q})$.

1.12.2 Rationality of local parameters at the cusps

We keep the assumptions and the notation of the previous section. Analytically, neighborhoods of E_{c_∞} were described in Sect. 1.4.1 with the aid of the parameters (z, u). Let $\widehat{S}$ denote the formal completion of $\bar{S}_{\mathcal{K}_N}$ along E. Let $r : \widehat{S} \to E$ be the infinitesimal retraction discussed in Sect. 1.10.1. If $i : E \hookrightarrow \widehat{S}$ is the closed embedding, then $r \circ i = Id_E$. If $\mathcal{I}$ is the sheaf of definition of E, then $\mathcal{N} = \mathcal{I}/\mathcal{I}^2$ is the conormal bundle to E and hence its analytification is the dual of the line bundle $\mathcal{T}$,

$$\mathcal{N}_{an} = \mathcal{T}^\vee. \tag{1.146}$$

Consider $r^*\mathcal{N}$ on $\widehat{S}$. The retraction allows us to split the exact sequence

$$0 \to \mathcal{N} \to i^*\Omega^1_{\widehat{S}} \to \Omega^1_E \to 0 \tag{1.147}$$

using $\Omega^1_E = i^*r^*\Omega^1_E \subset i^*\Omega^1_{\widehat{S}}$. Thus $i^*\Omega^1_{\widehat{S}} = \mathcal{N} \times \Omega^1_E$. The map $i \circ r : \widehat{S} \to \widehat{S}$ induces a sheaf homomorphism $r^*i^*\Omega^1_{\widehat{S}} \to \Omega^1_{\widehat{S}}$, which becomes the identity if we restrict it to E (i.e. follow it with i^*). By Nakayama's lemma, it is an isomorphism. It follows that

$$\Omega^1_{\widehat{S}} = r^*i^*\Omega^1_{\widehat{S}} = r^*\mathcal{N} \times r^*\Omega^1_E. \tag{1.148}$$

Let $x \in E$ and represent it by $u \in \mathbb{C}$ (modulo Λ). Then $q = e^{2\pi i z/M}$, where M is the width of the cusp (1.54), is a local analytic parameter on a classical neighborhood U_x of x which vanishes to first order along E. Note that q depends on the choice of u (see Remark below). It follows that $\mathrm{d}q$, the image of q in $\mathcal{I}_{an}/\mathcal{I}^2_{an}$, is a basis of $\mathcal{N}_{an}$ (on $U_x \cap E$). But

$$2\pi i \cdot \mathrm{d}z = M\frac{\mathrm{d}q}{q} \tag{1.149}$$

(mod $\langle \mathrm{d}u\rangle$) is independent of u [see (1.55)], so represents a global meromorphic section of $r^*\mathcal{N}_{an}$, with a simple pole along $E \subset \widehat{S}_\mathbb{C}$. By GAGA, this section is (meromorphic) algebraic.

Proposition 1.34 (i) *The section $2\pi i \cdot \mathrm{d}z$ mod $\langle \mathrm{d}u\rangle$ is $\mathcal{K}_N$-rational, i.e. it is the analytification of a section of $r^*\mathcal{N}$.* (ii) *The section $\Omega_B \cdot \mathrm{d}u$ is $\mathcal{K}_N$-rational, i.e. belongs to $H^0(E, \Omega^1_{E/\mathcal{K}_N})$.*

Proof The proof relies on the Kodaira–Spencer isomorphism $\mathrm{KS}(\Sigma)$ (1.141), which is a $\mathcal{K}_N$-rational (even $\mathcal{K}$-rational) algebraic isomorphism between $\mathcal{P} \otimes \mathcal{L}$ and Ω^1_S. As we have shown, it extends to a meromorphic homomorphism from $\mathcal{P} \otimes \mathcal{L}$ to $\Omega^1_{\bar{S}}$ over $\bar{S}$. Over $\widehat{S}$ it induces an isomorphism of $\mathcal{P}_0 \otimes \mathcal{L}$ onto $r^*\Omega^1_E \subset \Omega^1_{\widehat{S}}$ carrying the $\mathcal{K}_N$-rational section $\Omega_B \mathrm{d}\zeta_1 \otimes 2\pi i\mathrm{d}\zeta_3$ to $-\Omega_B\delta \cdot \mathrm{d}u$, proving part (ii) of the proposition. It also carries $2\pi i\mathrm{d}\zeta_2 \otimes 2\pi i\mathrm{d}\zeta_3$ to $2\pi i\mathrm{d}z$, but the latter is only meromorphic. We may summarize the situation over $\widehat{S}$ by the following commutative diagram with exact rows:

$$\begin{array}{ccccccccc} 0 \to & \widehat{\mathcal{I}} \otimes \mathcal{P}_0 \otimes \mathcal{L} & \to & \widehat{\mathcal{I}} \otimes \mathcal{P} \otimes \mathcal{L} & \to & \widehat{\mathcal{I}} \otimes \mathcal{P}_\mu \otimes \mathcal{L} & \to 0 \\ & \downarrow & & \downarrow \mathrm{KS}(\Sigma) & & \downarrow & \\ 0 \to & r^*\Omega^1_E & \to & \Omega^1_{\widehat{S}} & \to & r^*\mathcal{N} & \to 0 \end{array}. \tag{1.150}$$

Let h be a $\mathcal{K}_N$-rational local equation of E, i.e. a $\mathcal{K}_N$-rational section of $\mathcal{I}$ in some Zariski open U intersecting E non-trivially, vanishing to first order along $E \cap U$. The differential $\eta = h \cdot (2\pi i\mathrm{d}z)$ is regular on U, and to prove that it is $\mathcal{K}_N$-rational we may restrict it to $\widehat{S}$ and check rationality there. But in $\widehat{S}$ we have a $\mathcal{K}_N$-rational product decomposition $\Omega^1_{\widehat{S}} = r^*\mathcal{N} \times r^*\Omega^1_E$ and the projection of η to the second factor is 0, so it is enough to prove rationality of its projection to $r^*\mathcal{N}$. This projection is the image, under $\mathrm{KS}(\Sigma)$, of $h \cdot (2\pi i\mathrm{d}\zeta_2 \otimes 2\pi i\mathrm{d}\zeta_3 \mod \mathcal{P}_0 \otimes \mathcal{L})$, so our assertion follows from parts (i) and (ii) of the previous proposition. This proves that η, hence $h^{-1}\eta = 2\pi i\mathrm{d}z$ is a $\mathcal{K}_N$-rational differential. An alternative proof of part (ii) is to note that E is isogenous over $\mathcal{K}_N$ to B, so up to a $\mathcal{K}_N$-multiple has the same period.

Remark 1.1 The parameter q is not a well-defined parameter at x and depends not only on x, but also on the point u used to uniformize it. If we change u to $u+s$ ($s \in \Lambda$), then q is multiplied by the factor $e^{2\pi i\delta\bar{s}(u+s/2)/M}$, so although $\mathcal{O}^{hol}_{\bar{S}_{\mathbb{C}},x} \subset \widehat{\mathcal{O}}_{\bar{S}_{\mathbb{C}},x}$ and analytic parameters may be considered as formal parameters, the question whether q itself is $\mathcal{K}_N$-rational is not well defined (in sharp contrast to the case of modular curves!).

1.12.3 Normalizing the isomorphism det $\mathcal{P} \simeq \mathcal{L}$

Let us fix a nowhere vanishing section

$$\sigma \in H^0\left(S_{\mathcal{K}}, \det \mathcal{P} \otimes \mathcal{L}^{-1}\right). \tag{1.151}$$

This section is determined up to $\mathcal{K}^\times$. From now on, we shall use this section to identify $\det \mathcal{P}$ with $\mathcal{L}$ whenever such an identification is needed. From Corollary 1.33, we deduce that when we base change to $\mathbb{C}$, on each connected component X_Γ

$$\sigma \sim \Omega_B \cdot \sigma_{an}. \tag{1.152}$$

2 Picard modular schemes modulo an inert prime

2.1 The stratification

2.1.1 The three strata

Let p be a rational prime which is inert in $\mathcal{K}$ and relatively prime to $2N$. Then $\kappa_0 = R_0/pR_0$ is isomorphic to $\mathbb{F}_{p^2}$. We fix an algebraic closure κ of κ_0 and consider the characteristic p fiber

$$\bar{S}_\kappa = \bar{S} \times_{R_0} \kappa. \tag{2.1}$$

Unless otherwise specified, in this section we let S and $\bar{S}$ denote the characteristic p fibers S_κ and $\bar{S}_\kappa$. We also use the abbreviation $\omega_{\mathcal{A}}$ for $\omega_{\mathcal{A}/\bar{S}}$ etc.

Recall that an abelian variety over an algebraically closed field of characteristic p is called *supersingular* if the Newton polygon of its p-divisible group has a constant slope $1/2$. It is called *superspecial* if it is isomorphic to a product of supersingular elliptic curves. The following theorem combines various results proved in [4,39,40]. See also [8], Theorem 2.1.

Theorem 2.1 (i) *There exists a closed reduced 1-dimensional subscheme $S_{\mathrm{ss}} \subset \bar{S}$, disjoint from the cuspidal divisor (i.e. contained in S), which is uniquely characterized by the fact that for any geometric point x of S, the abelian variety $\mathcal{A}_x$ is supersingular if and only if x lies on S_{ss}. The scheme S_{ss} is defined over κ_0.*

(ii) *Let S_{ssp} be the singular locus of S_{ss}. Then x lies in S_{ssp} if and only if $\mathcal{A}_x$ is superspecial. If $x \in S_{\mathrm{ssp}}$, then*

$$\widehat{\mathcal{O}}_{S_{\mathrm{ss}},x} \simeq \kappa[[u,v]]/\left(u^{p+1}+v^{p+1}\right). \tag{2.2}$$

(iii) *Assume that N is large enough (depending on p). Then the irreducible components of S_{ss} are non-singular and in fact are all isomorphic to the Fermat curve $\mathcal{C}_p$ given by the equation*

$$x^{p+1}+y^{p+1}+z^{p+1}=0. \tag{2.3}$$

There are p^3+1 points of S_{ssp} on each irreducible component and through each such point pass $p+1$ irreducible components. Any two irreducible components are either disjoint or intersect transversally at a unique point.

(iv) *Without the assumption of N being large (but under $N \geq 3$ as usual), the irreducible components of S_{ss} may have multiple intersections with each other, including self-intersections. Their normalizations are nevertheless still isomorphic to $\mathcal{C}_p$.*

We call $\bar{S}_\mu = \bar{S} \backslash S_{\mathrm{ss}}$ (or $S_\mu = \bar{S}_\mu \cap S$) the *$\mu$-ordinary* or *generic* locus, $S_{\mathrm{gss}} = S_{\mathrm{ss}} \backslash S_{\mathrm{ssp}}$ the *general supersingular locus*, and S_{ssp} the *superspecial locus*. Then $\bar{S} = \bar{S}_\mu \cup S_{\mathrm{gss}} \cup S_{\mathrm{ssp}}$ is a stratification.

2.1.2 The p-divisible group

Let $x : \mathrm{Spec}(k) \to S$ (k an algebraically closed field) be a geometric point of S, $\mathcal{A}_x$ the corresponding fiber of $\mathcal{A}$, and $\mathcal{A}_x(p)$ its p-divisible group. Let $\mathfrak{G}$ be the p-divisible group of a supersingular elliptic curve over k (the group denoted by $G_{1,1}$ in the Manin-Dieudonné classification). The following theorem can be deduced from [4,39].

Theorem 2.2 (i) *If $x \in S_\mu$, then*

$$\mathcal{A}_x(p) \simeq (\mathcal{O}_{\mathcal{K}} \otimes \mu_{p^\infty}) \times \mathfrak{G} \times (\mathcal{O}_{\mathcal{K}} \otimes \mathbb{Q}_p/\mathbb{Z}_p). \tag{2.4}$$

(ii) *If $x \in S_{\mathrm{ss}}$, then $\mathcal{A}_x(p)$ is isogenous to $\mathfrak{G}^3$, and $x \in S_{\mathrm{ssp}}$ if and only if the two groups are isomorphic.*

While the p-divisible group of a μ-ordinary geometric fiber actually splits as a product of its multiplicative, local–local and étale parts, over the whole of S_μ we only get a filtration

$$0 \subset Fil^2\mathcal{A}(p) \subset Fil^1\mathcal{A}(p) \subset Fil^0\mathcal{A}(p) = \mathcal{A}(p) \tag{2.5}$$

by $\mathcal{O}_{\mathcal{K}}$-stable p-divisible groups. Here $gr^2 = Fil^2$ is of multiplicative type, $gr^1 = Fil^1/Fil^2$ is a local–local group and $gr^0 = Fil^0/Fil^1$ is étale, each of height 2 ($\mathcal{O}_{\mathcal{K}}$-height 1).

2.2 New relations between $\mathcal{P}$ and $\mathcal{L}$ in characteristic p

For proofs and more details on this subsection, see [8], Section 2.2.

2.2.1 The line bundles $\mathcal{P}_0$ and $\mathcal{P}_\mu$ over $\bar{S}_\mu$

Consider the universal semi-abelian variety $\mathcal{A}$ over the Zariski open set $\bar{S}_\mu$. Over the cuspidal divisor $C = \bar{S}\backslash S$, $\mathcal{P} = \omega_{\mathcal{A}}(\Sigma)$ admits a canonical filtration

$$0 \to \mathcal{P}_0 \to \mathcal{P} \to \mathcal{P}_\mu \to 0 \tag{2.6}$$

where $\mathcal{P}_0$ is the cotangent space to the abelian part of $\mathcal{A}$, and $\mathcal{P}_\mu$ is the Σ-component of the cotangent space to the toric part of $\mathcal{A}$. This filtration exists already in characteristic 0, but when we reduce the Picard surface modulo p, it extends to the whole of $\bar{S}_\mu$. Over the non-cuspidal part S_μ, we may set

$$\mathcal{P}_0 = \ker\left(\omega_{\mathcal{A}[p]^0} \to \omega_{\mathcal{A}[p]^\mu}\right) \tag{2.7}$$

where $\mathcal{A}[p]^\mu$ is the p-torsion in $\mathcal{A}(p)^\mu = Fil^2\mathcal{A}(p)$. Then $\mathcal{P}_\mu$ is identified with $\omega_{\mathcal{A}[p]^\mu}(\Sigma)$.

2.2.2 Frobenius and Verschiebung

Let $\mathcal{A}^{(p)} = \mathcal{A} \times_{\bar{S},\Phi} \bar{S}$ be the base change of $\mathcal{A}$ with respect to the absolute Frobenius morphism Φ of degree p of $\bar{S}$. The *relative Frobenius* is an $\mathcal{O}_{\bar{S}}$-linear isogeny $Frob_{\mathcal{A}}$: $\mathcal{A} \to \mathcal{A}^{(p)}$, characterized by the fact that $pr_1 \circ Frob_{\mathcal{A}}$ is the absolute Frobenius morphism of $\mathcal{A}$. Over S (but not over the boundary C), we have the dual abelian scheme $\mathcal{A}^t$, and the *Verschiebung* $Ver_{\mathcal{A}} : \mathcal{A}^{(p)} \to \mathcal{A}$ is the $\mathcal{O}_S$-linear isogeny which is dual to $Frob_{\mathcal{A}^t} : \mathcal{A}^t \to (\mathcal{A}^t)^{(p)}$.

We clearly have $\omega_{\mathcal{A}^{(p)}} = \omega_{\mathcal{A}}^{(p)}$, and we let

$$F : \omega_{\mathcal{A}}^{(p)} \to \omega_{\mathcal{A}}, \quad V : \omega_{\mathcal{A}} \to \omega_{\mathcal{A}}^{(p)} \tag{2.8}$$

be the $\mathcal{O}_{\bar{S}}$-linear maps of vector bundles induced by the isogenies $Frob_{\mathcal{A}}$ and $Ver_{\mathcal{A}}$ on the cotangent spaces. We refer to [8] for a discussion how to define V over the whole of $\bar{S}$, despite the fact that $Ver_{\mathcal{A}}$ is only defined over S.

Taking Σ-components, we get the map

$$V_{\mathcal{P}} : \mathcal{P} = \omega_{\mathcal{A}}(\Sigma) \to \omega_{\mathcal{A}}^{(p)}(\Sigma) = \omega_{\mathcal{A}}(\bar{\Sigma})^{(p)} = \mathcal{L}^{(p)}, \tag{2.9}$$

and taking the $\bar{\Sigma}$-component, we similarly get

$$V_{\mathcal{L}} : \mathcal{L} \to \mathcal{P}^{(p)}. \tag{2.10}$$

Proposition 2.3 *Over $\bar{S}_\mu$ both $V_{\mathcal{P}}$ and $V_{\mathcal{L}}$ are of rank 1,*

$$\mathcal{P}_0 = \ker V_{\mathcal{P}} \tag{2.11}$$

and the image of $V_{\mathcal{L}}$ is a direct sum complement to $\mathcal{P}_0^{(p)}$:

$$\mathcal{P}^{(p)} = \mathcal{P}_0^{(p)} \oplus V(\mathcal{L}). \tag{2.12}$$

Recall that over any base scheme in characteristic p, and for any *line* bundle $\mathcal{M}$, its base change $\mathcal{M}^{(p)}$ under the absolute Frobenius is canonically isomorphic to its pth power $\mathcal{M}^p$.

Corollary 2.4 *Over $\bar{S}_\mu$, $\mathcal{P}_\mu \simeq \mathcal{L}^p$, $\mathcal{P}_0 \simeq \mathcal{L}^{1-p}$, and $\mathcal{L}^{p^2} \simeq \mathcal{L}$. For $k \geq 1$ odd, $\mathcal{P}^{(p^k)} \simeq \mathcal{L}^{p-1} \oplus \mathcal{L}$. For $k \geq 2$ even, $\mathcal{P}^{(p^k)} \simeq \mathcal{L}^{1-p} \oplus \mathcal{L}^p$, but for $k = 0$ we only have an exact sequence*

$$0 \to \mathcal{L}^{1-p} \to \mathcal{P} \to \mathcal{L}^p \to 0. \tag{2.13}$$

Corollary 2.5 *Over $\bar{S}_\mu$, $\mathcal{L}^{p^2-1}$, $\mathcal{P}_\mu^{p^2-1}$ and $\mathcal{P}_0^{p+1}$ are trivial line bundles.*

2.2.3 Extending the filtration on $\mathcal{P}$ over S_{gss}

In order to determine to what extent the filtration on $\mathcal{P}$ and the relation between $\mathcal{L}$ and the two graded pieces of the filtration extend into the supersingular locus, we have to employ Dieudonné theory. The following is proved in [8].

Proposition 2.6 (i) *Let $\mathcal{P}_0 = \ker V_{\mathcal{P}}$ (this agrees with what was denoted by $\mathcal{P}_0$ over $\bar{S}_\mu$). Then outside S_{ssp}, $V(\mathcal{P}) = \mathcal{L}^{(p)}$ and $\mathcal{P}_0$ is a rank 1 submodule.*
(ii) *Let $\mathcal{P}_\mu = \mathcal{P}/\mathcal{P}_0$. Then outside S_{ssp} we have $\mathcal{P}_\mu \simeq \mathcal{L}^p$ and $\mathcal{P}_0 \simeq \mathcal{L}^{1-p}$.*

For $V_{\mathcal{L}}$, we similarly get the following.

Proposition 2.7 *Outside S_{ssp}, $V_{\mathcal{L}}$ maps $\mathcal{L}$ injectively onto a sub-line-bundle of $\mathcal{P}^{(p)}$.*

At a superspecial point, both $V_{\mathcal{P}}$ and $V_{\mathcal{L}}$ vanish.

2.2.4 The Hasse invariant

As we have just seen, the fact that $V_{\mathcal{P}}$ and $V_{\mathcal{L}}$ are both of rank 1 "extends" across the general supersingular locus S_{gss}. However, while $\mathrm{Im}(V_{\mathcal{L}})$ and $\ker(V_{\mathcal{P}}^{(p)}) = \mathcal{P}_0^{(p)}$ made up a frame of $\mathcal{P}$ over $\bar{S}_\mu$, over S_{gss} these two line bundles coincide. To state a more precise result, we make the following definition.

Definition 2.8 The *Hasse invariant* is

$$h_{\bar{\Sigma}} = V_{\mathcal{P}}^{(p)} \circ V_{\mathcal{L}} \in \mathrm{Hom}\left(\mathcal{L}, \mathcal{L}^{(p^2)}\right). \tag{2.14}$$

As $\mathcal{L}^{(p^2)} \simeq \mathcal{L}^{p^2}$, the Hasse invariant is a global section of $\mathcal{L}^{p^2-1}$, i.e. a modular form of weight $p^2 - 1$ over κ,

$$h_{\bar{\Sigma}} \in M_{p^2-1}(N, \kappa). \tag{2.15}$$

It turns out that $h_{\bar{\Sigma}}$ has a *simple* zero along the supersingular locus S_{ss}. Once again, this requires a little computation with Dieudonné modules. Equivalently, we have the following theorem.

Theorem 2.9 *The divisor of $h_{\bar{\Sigma}}$ is S_{ss} (with its reduced subscheme structure).*

2.3 The open Igusa surfaces

2.3.1 The Igusa scheme

Let $N \geq 3$ as always, and let $\mathcal{M}$ be the moduli problem of Sect. 1.3.1. Let $n \geq 1$ and consider the following moduli problem on κ_0-algebras:

- $\mathcal{M}_{Ig(p^n)}(R)$ is the set of isomorphism classes of pairs $(\underline{A}, \varepsilon)$ where $\underline{A} \in \mathcal{M}(R)$ and

$$\varepsilon : \delta_{\mathcal{K}}^{-1}\mathcal{O}_{\mathcal{K}} \otimes \mu_{p^n} \hookrightarrow A[p^n] \tag{2.16}$$

is a closed immersion of $\mathcal{O}_{\mathcal{K}}$-group schemes over R.

It is clear that if $(\underline{A}, \varepsilon) \in \mathcal{M}_{Ig(p^n)}(R)$, then A is fiber-by-fiber μ-ordinary and therefore $\underline{A} \in \mathcal{M}(R)$ defines an R-point of S_μ. The image of ε is then $A[p^n]^\mu$, the maximal subgroup scheme of $A[p^n]$ of multiplicative type. It is also clear that the functor $R \rightsquigarrow \mathcal{M}_{Ig(p^n)}(R)$ is relatively representable over $\mathcal{M}$, and therefore as $N \geq 3$ and $\mathcal{M}$ is representable, this functor is also representable by a scheme $Ig_\mu(p^n)$ which maps to S_μ. See [23] for the notion of relative representability. We call $Ig_\mu(p^n)$ the *Igusa scheme of level* p^n.

Proposition 2.10 *The morphism $\tau : Ig_\mu(p^n) \to S_\mu$ is finite and étale, with the Galois group $\Delta(p^n) = (\mathcal{O}_\mathcal{K}/p^n\mathcal{O}_\mathcal{K})^\times$ acting as a group of deck transformations.*

Proof Every μ-ordinary abelian variety has a *unique* finite flat $\mathcal{O}_\mathcal{K}$-subgroup scheme of multiplicative type $A[p^n]^\mu$ of rank p^{2n}. Such a subgroup scheme is, locally in the étale topology, isomorphic to $\delta_\mathcal{K}^{-1}\mathcal{O}_\mathcal{K} \otimes \mu_{p^n}$, and any two isomorphisms differ by a unique automorphism of $\delta_\mathcal{K}^{-1}\mathcal{O}_\mathcal{K} \otimes \mu_{p^n}$. But $\Delta(p^n) = Aut_{\mathcal{O}_\mathcal{K}}(\delta_\mathcal{K}^{-1}\mathcal{O}_\mathcal{K} \otimes \mu_{p^n})$. If we let $\gamma \in \Delta(p^n)$ act on the pair $(\underline{A}, \varepsilon)$ via

$$\gamma((\underline{A}, \varepsilon)) = \left(\underline{A}, \varepsilon \circ \gamma^{-1}\right) \tag{2.17}$$

$\Delta(p^n)$ becomes a group of deck transformation and the proof is complete.

2.3.2 A compactification over the cusps

The proof of the following proposition mimics the construction of $\bar{S}$. We omit it.

Proposition 2.11 *Let $\overline{Ig}_\mu(p^n)$ be the normalization of $\bar{S}_\mu = \bar{S}\backslash S_{ss}$ in $Ig_\mu(p^n)$. Then, $\overline{Ig}_\mu(p^n) \to \bar{S}_\mu$ is finite étale and the action of $\Delta(p^n)$ extends to it. The boundary $\overline{Ig}_\mu(p^n)\backslash Ig_\mu(p^n)$ is non-canonically identified with $\Delta(p^n) \times C$.*

We define similarly Ig_μ^*, and note that it is finite étale over S_μ^*.

Proposition 2.12 *Let $\mathcal{A}$ denote the pull-back of the universal semi-abelian variety from $\bar{S}_\mu$ to $\overline{Ig}_\mu(p^n)$. Then $\mathcal{A}$ is equipped with a canonical Igusa level structure*

$$\varepsilon : \delta_\mathcal{K}^{-1}\mathcal{O}_\mathcal{K} \otimes \mu_{p^n} \simeq \mathcal{A}[p^n]^\mu. \tag{2.18}$$

Over C and after base change to R_N/pR_N the toric part of $\mathcal{A}$ is locally Zariski of the form $\mathfrak{a} \otimes \mathbb{G}_m$ and ε is then an $\mathcal{O}_\mathcal{K}$-linear isomorphism between $\delta_\mathcal{K}^{-1}\mathcal{O}_\mathcal{K} \otimes \mu_{p^n}$ and $\mathfrak{a} \otimes \mu_{p^n}$.

2.3.3 A trivialization of $\mathcal{L}$ over the Igusa surface

From now on, we focus on $\overline{Ig}_\mu = \overline{Ig}_\mu(p)$ although similar results hold when $n > 1$, and would be instrumental in the study of p-adic modular forms. The vector bundle $\omega_\mathcal{A}$ pulls back to a similar vector bundle over $\overline{Ig}_\mu$. But there

$$\omega_\mathcal{A}^\mu := \omega_{\mathcal{A}[p]^\mu} \tag{2.19}$$

is a rank 2 quotient bundle stable under $\mathcal{O}_\mathcal{K}$ (of type $(1, 1)$), and the isomorphism ε induces an isomorphism

$$\varepsilon^* : \omega_\mathcal{A}^\mu \simeq \omega_{\delta_\mathcal{K}^{-1}\mathcal{O}_\mathcal{K}\otimes\mu_p}. \tag{2.20}$$

Now $Lie(\delta_\mathcal{K}^{-1}\mathcal{O}_\mathcal{K} \otimes \mu_p) = \delta_\mathcal{K}^{-1}\mathcal{O}_\mathcal{K} \otimes Lie(\mu_p) = \delta_\mathcal{K}^{-1}\mathcal{O}_\mathcal{K} \otimes Lie(\mathbb{G}_m)$ and by duality

$$\omega_{\delta_\mathcal{K}^{-1}\mathcal{O}_\mathcal{K}\otimes\mu_p} = \mathcal{O}_\mathcal{K} \otimes \omega_{\mathbb{G}_m}, \tag{2.21}$$

with $1 \otimes dT/T$ as a generator (if T is the parameter of $\mathbb{G}_m$). Here we have used the fact that the $\mathbb{Z}$-dual of $\delta_{\mathcal{K}}^{-1}\mathcal{O}_{\mathcal{K}}$ is $\mathcal{O}_{\mathcal{K}}$ via the trace pairing. This is the constant vector bundle $\mathcal{O}_{\mathcal{K}} \otimes R = R(\Sigma) \oplus R(\bar{\Sigma})$.

Proposition 2.13 *The line bundles $\mathcal{L}$, $\mathcal{P}_0$ and $\mathcal{P}_\mu$ are trivial over $\overline{Ig}_\mu$.*

Proof Use ε^* as an isomorphism between vector bundles and note that $\mathcal{L} = \omega_{\mathcal{A}}^{\mu}(\bar{\Sigma})$ and $\mathcal{P}_\mu = \omega_{\mathcal{A}}^{\mu}(\Sigma)$. The relation $\mathcal{P}_0 \otimes \mathcal{P}_\mu = \det \mathcal{P} \simeq \mathcal{L}$ implies the triviality of $\mathcal{P}_0$ as well.

Note that the trivialization of $\mathcal{L}$ and $\mathcal{P}_\mu$ is canonical, because it uses only the tautological map ε which exists over the Igusa scheme. The trivialization of $\mathcal{P}_0$ on the other hand depends on how we realize the isomorphism $\det \mathcal{P} \simeq \mathcal{L}$.

We can now give an alternative proof to the fact that $\mathcal{L}^{p^2-1}$ and $\mathcal{P}_\mu^{p^2-1}$ are trivial on $\bar{S}_\mu$. Denote by $\mathcal{O}_{Ig}$ the structure sheaf of $\overline{Ig}_\mu$. By the projection formula, $\tau_*(\tau^*\mathcal{L}) \simeq \mathcal{L} \otimes \tau_*\mathcal{O}_{Ig}$. Taking determinants, we get

$$\det \tau_*(\tau^*\mathcal{L}) \simeq \mathcal{L}^{p^2-1} \otimes \det \tau_*\mathcal{O}_{Ig}. \tag{2.22}$$

As $\tau^*\mathcal{L} \simeq \mathcal{O}_{Ig}$, we get that $\mathcal{L}^{p^2-1} \simeq \mathcal{O}_{\bar{S}}$. The same argument works for $\mathcal{P}_\mu$ and for $\mathcal{P}_0$. The fact that $\mathcal{P}_0^{p+1}$ is already trivial could be deduced by a similar argument had we worked out an analogue of $Ig(p)$ classifying *symplectic isomorphisms* of $\mathfrak{G}[p]$ with $gr^1A[p]$. The role of $\Delta(p)$ for such a moduli space would be assumed by

$$\Delta^1(p) = \ker(N : (\mathcal{O}_{\mathcal{K}}/p\mathcal{O}_{\mathcal{K}})^\times \to \mathbb{F}_p^\times), \tag{2.23}$$

which is a group of order $p+1$. We do not go any further in this direction here.

2.4 Compactification of the Igusa surface along the supersingular locus

2.4.1 Extracting a p^2-1 root from $h_{\bar{\Sigma}}$ over $\overline{Ig}_\mu$

Let a be the canonical nowhere vanishing section of $\mathcal{L}$ over $\overline{Ig}_\mu$ which is sent to $e_{\bar{\Sigma}} \cdot (1 \otimes dT/T)$ under the trivialization

$$\varepsilon^* : \mathcal{L} = \omega_{\mathcal{A}}^{\mu}(\bar{\Sigma}) \simeq (\mathcal{O}_{\mathcal{K}} \otimes \omega_{\mathbb{G}_m})(\bar{\Sigma}) = R(\bar{\Sigma}). \tag{2.24}$$

Here R is any R_0/pR_0-algebra over which we choose to work. In other words, $a = (\varepsilon^*)^{-1}(e_{\bar{\Sigma}} \cdot 1 \otimes dT/T)$. Dually, a is the homomorphism from $Lie(\mathcal{A})(\bar{\Sigma})$ to $\delta_{\mathcal{K}}^{-1} \otimes Lie(\mathbb{G}_m)(\bar{\Sigma})$ arising from ε^{-1}. Let $a(k) = a^{\otimes k} \in H^0(\overline{Ig}_\mu, \mathcal{L}^k)$.

Proposition 2.14 (i) *Let $\gamma \in \Delta(p) = (\mathcal{O}_{\mathcal{K}}/p\mathcal{O}_{\mathcal{K}})^\times$. Then $\Delta(p)$ acts on $H^0(\overline{Ig}_\mu, \mathcal{L})$ and*

$$\gamma^* a = \bar{\Sigma}(\gamma)^{-1} \cdot a. \tag{2.25}$$

(ii) *The section a is a p^2-1 root of the Hasse invariant over $\overline{Ig}_\mu$, i.e.*

$$a\left(p^2-1\right) = h_{\bar{\Sigma}}. \tag{2.26}$$

Proof (i) This part is a restatement of the action of $\Delta(p)$. At two points of $Ig_\mu(R)$ lying over the same point of $S_\mu(R)$ and differing by the action of $\gamma \in \Delta(p)$, the canonical embeddings

$$\delta_{\mathcal{K}}^{-1} \otimes \mu_p \hookrightarrow A[p] \tag{2.27}$$

differ by $\iota(\gamma)$ (2.17). The induced trivializations of $Lie(A)(\bar{\Sigma})$ differ by $\bar{\Sigma}(\gamma)$ and by duality we get (i).

(ii) Since over any $\mathbb{F}_p$-base, $Ver_{\mathbb{G}_m} = 1$, we have a commutative diagram

$$\begin{array}{lll} Lie(\mathcal{A})(\bar{\Sigma})^{(p^2)} & \overset{V_*^2}{\to} Lie(\mathcal{A})(\bar{\Sigma}) & \\ \downarrow a^{(p^2)} & \downarrow a & . \\ \delta_{\mathcal{K}}^{-1} \otimes Lie(\mathbb{G}_m)(\bar{\Sigma}) = & \delta_{\mathcal{K}}^{-1} \otimes Lie(\mathbb{G}_m)(\bar{\Sigma}) & \end{array} \tag{2.28}$$

Using the isomorphism $Lie(\mathcal{A})(\bar{\Sigma})^{(p^2)} \simeq Lie(\mathcal{A})(\bar{\Sigma})^{p^2}$, we get the commutative diagram

$$\begin{array}{lll} Lie(\mathcal{A})(\bar{\Sigma})^{p^2} & \overset{h_{\bar{\Sigma}}}{\to} Lie(\mathcal{A})(\bar{\Sigma}) & \\ \downarrow a(p^2) & \downarrow a & , \\ \delta_{\mathcal{K}}^{-1} \otimes Lie(\mathbb{G}_m)(\bar{\Sigma}) = & \delta_{\mathcal{K}}^{-1} \otimes Lie(\mathbb{G}_m)(\bar{\Sigma}) & \end{array} \tag{2.29}$$

from which we deduce that $h_{\bar{\Sigma}} = a(p^2 - 1)$.

2.4.2 The compactification $\overline{Ig}$ of $\overline{Ig}_\mu$

In this section, we follow the method outlined in [1, Sections 6-9] and [12] for Hilbert modular varieties. Quite generally, let $L \to X$ be a line bundle associated with an invertible sheaf $\mathcal{L}$ on a scheme X. Write L^n for the line bundle $L^{\otimes n}$ over X. Let $s : X \to L^n$ be a section. Consider the fiber product

$$Y = L \times_{L^n} X \tag{2.30}$$

where the two maps to L^n are $\lambda \mapsto \lambda^n$ and s. Let $p : Y \overset{pr_2}{\to} X$ be the projection which factors also as $Y \overset{pr_1}{\to} L \to L^n \to X$ (since $X \overset{s}{\to} L^n \to X$ is the identity). Consider

$$p^*L = L \times_X (L \times_{L^n} X). \tag{2.31}$$

This line bundle on Y has a tautological section $t : Y \to p^*L$,

$$t : y = (\lambda, x) \mapsto (\lambda, y) = (\lambda, (\lambda, x)) \tag{2.32}$$

Here $s(x) = \lambda^n$ and

$$t^n(y) = (\lambda^n, y) = (s(x), y) = p^*s(y) \tag{2.33}$$

so t is an nth root of p^*s. Moreover, Y has the *universal property* with respect to extracting nth roots from s: If $p_1 : Y_1 \to X$, and $t_1 \in \Gamma(Y_1, p_1^*L)$ is such that $t_1^n = p_1^*s$, then there exists a unique morphism $h : Y_1 \to Y$ covering the two maps to X such that $t_1 = h^*t$.

The map $L \to L^n$ is *finite flat of degree n* and if n is invertible on the base, finite étale is away from the zero section. Indeed, locally on X it is the map $\mathbb{A}^1 \times X \to \mathbb{A}^1 \times X$ which is just raising to nth power in the first coordinate. By base change, it follows that the same is true for the map $p : Y \to X$: this map is finite flat of degree n and étale away from the vanishing locus of the section s (assuming n is invertible). We remark that if L is the trivial line bundle, we recover usual Kummer theory.

Applying this in our example with $n = p^2 - 1$, we define the *complete Igusa surface of level p*, $\overline{Ig} = \overline{Ig}(p)$ as

$$\overline{Ig} = \mathcal{L} \times_{\mathcal{L}^{p^2-1}} \bar{S} \tag{2.34}$$

where the map $\overline{S} \to \mathcal{L}^{p^2-1}$ is $h_{\bar{\Sigma}}$. From the universal property and part (ii) of Proposition 2.14 we get a map of $\bar{S}$-schemes

$$\overline{Ig}_\mu \to \overline{Ig}. \tag{2.35}$$

This map is an isomorphism over $\bar{S}_\mu$ because both schemes are étale torsors for $\Delta(p) = (\mathcal{O}_\mathcal{K}/p\mathcal{O}_\mathcal{K})^\times$ and the map respects the action of this group. We summarize the discussion in the following theorem (for the last point, consult [32], Proposition 2, p.198).

Theorem 2.15 *The morphism $\tau : \overline{Ig} \to \bar{S}$ satisfies the following properties:*

(i) *It is finite flat of degree $p^2 - 1$, étale over $\bar{S}_\mu$, totally ramified over S_{ss}.*
(ii) *$\Delta(p)$ acts on $\overline{Ig}$ as a group of deck transformations and the quotient is $\bar{S}$.*
(iii) *Let $s_0 \in S_{\mathrm{gss}}(\bar{\mathbb{F}}_p)$. Then there exist local parameters u, v at s_0 such that $\widehat{\mathcal{O}}_{S,s_0} = \bar{\mathbb{F}}_p[[u,v]]$, $S_{\mathrm{gss}} \subset S$ is formally defined by $u = 0$, and if $\tilde{s}_0 \in Ig$ maps to s_0 under τ, then $\widehat{\mathcal{O}}_{Ig,\tilde{s}_0} = \bar{\mathbb{F}}_p[[w,v]]$ where $w^{p^2-1} = u$. In particular, Ig is regular in codimension 1.*
(iv) *Let $s_0 \in S_{\mathrm{ssp}}(\bar{\mathbb{F}}_p)$. Then there exist local parameters u, v at s_0 such that $\widehat{\mathcal{O}}_{S,s_0} = \bar{\mathbb{F}}_p[[u,v]]$, $S_{\mathrm{ss}} \subset S$ is formally defined at s_0 by $u^{p+1} + v^{p+1} = 0$, and if $\tilde{s}_0 \in Ig$ maps to s_0 under τ, then*

$$\widehat{\mathcal{O}}_{Ig,\tilde{s}_0} = \bar{\mathbb{F}}_p[[w,u,v]]/\left(w^{p^2-1} - u^{p+1} - v^{p+1}\right) \tag{2.36}$$

In particular, $\tilde{s}_0$ is a normal singularity of Ig.

2.4.3 Irreducibility of Ig

So far we have avoided the delicate question of whether $\overline{Ig}$ is "relatively irreducible", i.e. whether $\tau^{-1}(T)$ is irreducible if $T \subset \bar{S}$ is an irreducible (equivalently, connected) component. Using an idea of Katz, and following the approach taken by Ribet [33], the irreducibility of $\tau^{-1}(T)$ could be proven for any level p^n if we could prove the following:

- Let $q = p^2$. For any r sufficiently large and for any $\gamma \in (\mathcal{O}_\mathcal{K}/p^n\mathcal{O}_\mathcal{K})^\times$ there exists a μ-ordinary abelian variety with PEL structure $\underline{A} \in S_\mu(\mathbb{F}_{q^r})$ such that the image of $Gal(\bar{\mathbb{F}}_q/\mathbb{F}_{q^r})$ in

$$Aut\left(Isom_{\bar{\mathbb{F}}_q}(\delta_\mathcal{K}^{-1} \otimes \mu_{p^n}, A[p^n]^\mu)\right) = (\mathcal{O}_\mathcal{K}/p^n\mathcal{O}_\mathcal{K})^\times \tag{2.37}$$

contains γ.

See also the discussion in 5.2.5. Instead, we shall give a different argument valid for the case $n = 1$.

Proposition 2.16 *The morphism $\tau : \overline{Ig} \to \bar{S}$ induces a bijection on irreducible components.*

Proof Since $\overline{Ig}$ is a normal surface, connected components and irreducible components are the same. Let T be a connected component of $\bar{S}$ and $T_{\mathrm{ss}} = T \cap S_{\mathrm{ss}}$. Let $\tau^{-1}(T) = \coprod Y_i$ be the decomposition into connected components. As τ is finite and flat, each $\tau(Y_i) = T$. Since τ is totally ramified over T_{ss}, there is only one Y_i.

3 Modular forms modulo *p* and the theta operator

3.1 Modular forms mod *p* as functions on *Ig*

3.1.1 Representing modular forms by functions on Ig

The Galois group $\Delta(p) = (\mathcal{O}_\mathcal{K}/p\mathcal{O}_\mathcal{K})^\times$ acts on the coordinate ring $H^0(Ig_\mu, \mathcal{O})$ and we let $H^0(Ig_\mu, \mathcal{O})^{(k)}$ be the subspace where it acts via the character $\bar{\Sigma}^k$. Then

$$H^0(Ig_\mu, \mathcal{O}) = \bigoplus_{k=0}^{p^2-2} H^0(Ig_\mu, \mathcal{O})^{(k)} \tag{3.1}$$

and each $H^0(Ig_\mu, \mathcal{O})^{(k)}$ is free of rank 1 over $H^0(S_\mu, \mathcal{O}) = H^0(Ig_\mu, \mathcal{O})^{(0)}$.

For any $0 \le k$, the map $f \mapsto f/a(k)$ is an embedding

$$M_k(N, \kappa_0) \hookrightarrow H^0(Ig_\mu, \mathcal{O})^{(k)}. \tag{3.2}$$

Lemma 3.1 *Fix $0 \le k < p^2 - 1$. Then we have a surjective homomorphism*

$$\bigoplus_{n\ge 0} M_{k+n(p^2-1)}(N, \kappa_0) \twoheadrightarrow H^0(Ig_\mu, \mathcal{O})^{(k)}. \tag{3.3}$$

Proof Take $f \in H^0(Ig_\mu, \mathcal{O})^{(k)}$, so that $f \cdot a(k) \in H^0(Ig_\mu, \mathcal{L}^k)^{(0)}$ and hence descends to $g \in H^0(S_\mu, \mathcal{L}^k)$. This g may have poles along S_{ss}, but some $h_{\bar\Sigma}^n g$ will extend holomorphically to S and hence represents a modular form of weight $k + n(p^2 - 1)$, which will map to f because $a(k + n(p^2 - 1)) = h_{\bar\Sigma}^n a(k)$.

Proposition 3.2 *The resulting ring homomorphism*

$$r : \bigoplus_{k\ge 0} M_k(N, \kappa_0) \twoheadrightarrow H^0(Ig_\mu, \mathcal{O}) \tag{3.4}$$

obtained by dividing a modular form of weight k by $a(k)$ is surjective, respects the $\mathbb{Z}/(p^2-1)$ $\mathbb{Z}$-grading on both sides, and its kernel is the ideal generated by $(h_{\bar\Sigma} - 1)$.

Proof We only have to prove that anything in $\ker(r)$ is a multiple of $h_{\bar\Sigma} - 1$, the rest being clear. Since r respects the grading, we may assume that for some $k \ge 0$ we have $f_j \in M_{k+j(p^2-1)}(S, \kappa_0)$ and $f = \sum_{j=0}^m f_j \in \ker(r)$, i.e.

$$\sum_{j=0}^{m} a(k)^{-1} h_{\bar\Sigma}^{-j} f_j = 0. \tag{3.5}$$

But then $f_m = -h_{\bar\Sigma}^m \left(\sum_{j=0}^{m-1} h_{\bar\Sigma}^{-j} f_j\right)$, so $\sum_{j=0}^m f_j = \sum_{j=0}^{m-1}(1 - h_{\bar\Sigma}^{m-j}) f_j$ belongs to $(1 - h_{\bar\Sigma})$.

As a result, we get that

$$Ig_\mu^* = \mathrm{Spec}\left(\bigoplus_{k\ge 0} M_k(N, \kappa_0)/(h_{\bar\Sigma} - 1)\right) \tag{3.6}$$

and

$$S_\mu^* = \mathrm{Spec}\left(\bigoplus_{k\ge 0} M_{k(p^2-1)}(N, \kappa_0)/(h_{\bar\Sigma} - 1)\right). \tag{3.7}$$

3.1.2 Fourier–Jacobi expansions modulo p

The arithmetic Fourier–Jacobi expansion (1.125) depended on a choice of a nowhere vanishing section s of $\mathcal{L}$ along the boundary $C = \bar S \backslash S$ of $\bar S$. As the boundary $\tilde C = \overline{Ig}_\mu \backslash Ig_\mu$ is (non-canonically) identified with $\Delta(p) \times C$, we may "compute" the Fourier–Jacobi expansion on the Igusa surface rather than on S. But on the Igusa surface, a is a canonical choice for such an s. We may therefore associate a *canonical* Fourier–Jacobi expansion

$$\widetilde{FJ}(f)=\sum_{m=0}^{\infty}c_m(f)\in\prod_{m=0}^{\infty}H^0(\tilde{C},\mathcal{N}^m) \tag{3.8}$$

along the boundary of Ig, with every

$$f\in M_*(N,R)=\bigoplus_{k=0}^{\infty}M_k(N,R) \tag{3.9}$$

(R a κ_0-algebra). The following proposition becomes almost a *tautology*.

Proposition 3.3 *The Fourier–Jacobi expansion $\widetilde{FJ}(h_{\bar{\Sigma}})$ of the Hasse invariant is 1. Moreover, for f_1 and f_2 in the graded ring $M_*(N,R)$, $r(f_1)=r(f_2)$ if and only if $\widetilde{FJ}(f_1)=\widetilde{FJ}(f_2)$.*

Proof The first statement is tautologically true. For the second, note that for $f\in M_k(N,R)$, $\widetilde{FJ}(f)$ is the (expansion of the) image of $f/a(k)$ in $H^0(\tilde{C},\mathcal{O}_{\widehat{Ig}})$ where $\widehat{Ig}$ is the formal completion of Ig along $\tilde{C}$, while $r(f)$ is the image of $f/a(k)$ in $H^0(\overline{Ig}_\mu,\mathcal{O})$. The proposition follows from the fact that by Proposition 2.16 the irreducible components of $\overline{Ig}_\mu$ are in bijection with the connected components of $\bar{S}$, so every irreducible component of $\overline{Ig}_\mu$ contains at least one cuspidal component ("q-expansion principle"). A function on $\overline{Ig}_\mu$ that vanishes in the formal neighborhood of any cuspidal component must therefore vanish on any irreducible component, so is identically 0.

3.1.3 The filtration of a modular form modulo p

Let $f\in M_k(N,R)$, where R is a κ_0-algebra as before. Define the *filtration* $\omega(f)$ to be the minimal $j\geq 0$ such that $r(f)=r(f')$ (equivalently $FJ(f)=FJ(f'))$ for some $f'\in M_j(N,R)$. The following proposition follows immediately from previous results.

Proposition 3.4 *Let $f\in M_k(N,R)$. Then $0\leq\omega(f)\leq k$ and*

$$\omega(f)\equiv k \mod (p^2-1). \tag{3.10}$$

Let $\omega(f)=k-(p^2-1)n$. Then n is the order of vanishing of f along S_{ss}. Equivalently, $k-\omega(f)$ is the order of vanishing of the pull-back of f to Ig along Ig_{ss}. In addition, $\omega(f^m)=m\omega(f)$.

3.2 The theta operator

3.2.1 Definition of $\Theta(f)$

We work over $\kappa=\bar{\mathbb{F}}_p$. Let S be the (open) Picard surface over κ and $Ig=Ig(p)$ the Igusa surface of level p (completed along the supersingular locus as explained above). To simplify the notation, we denote by $Z=S_{\mathrm{ss}}=S\backslash S_\mu$ the supersingular locus of S, by $\tilde{Z}=Ig_{\mathrm{ss}}=Ig\backslash Ig_\mu$ its pre-image under the covering map $\tau:Ig\to S$, by $Z'=S_{\mathrm{gss}}=S_{\mathrm{ss}}\backslash S_{\mathrm{ssp}}$ the smooth part of Z, and by $\tilde{Z}'=Ig_{\mathrm{gss}}=Ig_{\mathrm{ss}}\backslash Ig_{\mathrm{ssp}}$ the pre-image of Z' under τ.

Let $f\in H^0(S,\mathcal{L}^k)$. Then $\tau^*f/a^k\in H^0(Ig_\mu,\mathcal{O})$ has a pole of order at most k along $\tilde{Z}$, and the Galois group acts on it via $\bar{\Sigma}^k$. Let

$$\eta_f=d(\tau^*f/a^k)\in H^0(Ig_\mu,\Omega^1_{Ig})=H^0(Ig_\mu,\tau^*\Omega^1_S). \tag{3.11}$$

The Kodaira–Spencer isomorphism $\mathrm{KS}(\Sigma)$ is an isomorphism

$$\mathrm{KS}(\Sigma):\mathcal{P}\otimes\mathcal{L}\simeq\Omega^1_S. \tag{3.12}$$

Let

$$\psi=(V_{\mathcal{P}}\otimes 1)\circ\mathrm{KS}(\Sigma)^{-1}:\Omega^1_S\to\mathcal{L}^{(p)}\otimes\mathcal{L}\simeq\mathcal{L}^{p+1}. \tag{3.13}$$

We denote by ψ also the map induced on the base change of these vector bundles by τ^* to Ig and consider $\psi(\eta_f)$. As $\Delta(p)$ still acts on $\psi(\eta_f)$ via $\bar{\Sigma}^k$, its action on $a^k\psi(\eta_f)$ is trivial, so this section descends to S_μ. We define

$$\Theta(f) = a^k\psi(\eta_f) \in H^0(S_\mu, \mathcal{L}^{k+p+1}). \tag{3.14}$$

A priori, this extends only to a meromorphic modular form of weight $k+p+1$, as it may have poles along Z.

3.2.2 The main theorem

For the formulation of the next theorem, we need to define what we mean by the *standard cuspidal component* of $\bar{S}$ or $\overline{Ig}$. Since its definition involves a transition back and forth between $\mathbb{C}$ and κ, we need to fix, besides the embedding of R_N in $\mathbb{C}$ also a homomorphism

$$R_N \to \kappa \tag{3.15}$$

extending the map $R_0 \to \kappa_0 \subset \kappa$, and we let $\mathfrak{P}$ be its kernel (a prime above p).

Recall that according to [2,28] the cuspidal scheme $C = \bar{S}\backslash S$ classifies $\mathcal{O}_\mathcal{K}$-semi-abelian varieties with level N structure. The *standard component* of C over $\mathbb{C}$ is the component which classifies extensions of the elliptic curve $\mathbb{C}/\mathcal{O}_\mathcal{K}$ by the $\mathcal{O}_\mathcal{K}$-torus $\mathcal{O}_\mathcal{K} \otimes \mathbb{C}^\times$ (thus sits over a cusp of type $(\mathcal{O}_\mathcal{K}, \mathcal{O}_\mathcal{K})$ in $S^*_\mathbb{C}$), together with a level-N structure (α, β, γ) (see [2], I.4.2 and Sect. 1.6.2), where

$$\alpha : \mathcal{O}_\mathcal{K}/N\mathcal{O}_\mathcal{K} = \mathcal{O}_\mathcal{K} \otimes \mathbb{Z}/N\mathbb{Z} \to \mathcal{O}_\mathcal{K} \otimes \mathbb{C}^\times \tag{3.16}$$

is given by $1 \otimes (a \mapsto \exp(2\pi ia/N))$ and

$$\beta : \mathcal{O}_\mathcal{K}/N\mathcal{O}_\mathcal{K} = N^{-1}\mathcal{O}_\mathcal{K}/\mathcal{O}_\mathcal{K} \to \mathbb{C}/\mathcal{O}_\mathcal{K} \tag{3.17}$$

is the canonical embedding. (The splitting γ varies along the component.) The standard component of C over R_N is the one which becomes this component after base change to $\mathbb{C}$. The standard component of C over κ is the reduction modulo $\mathfrak{P}$ of the standard component of C over R_N. Finally, $\overline{Ig}$ maps to $\bar{S}$ (over κ) and the cuspidal components mapping to a given component E of C are classified by the embedding of $\delta_\mathcal{K}^{-1}\mathcal{O}_\mathcal{K} \otimes \mu_p$ in the toric part of $\mathcal{A}$. Since the toric part of the universal semi-abelian variety over the standard component is $\mathcal{O}_\mathcal{K} \otimes \mathbb{G}_m$, we may define the standard cuspidal component of $\overline{Ig}$ to be the component where the map

$$\varepsilon : \delta_\mathcal{K}^{-1}\mathcal{O}_\mathcal{K} \otimes \mu_p \to \mathcal{O}_\mathcal{K} \otimes \mathbb{G}_m \tag{3.18}$$

is the natural embedding. Here we use the fact that

$$\delta_\mathcal{K}^{-1}\mathcal{O}_\mathcal{K} \otimes \mu_p = \mathcal{O}_\mathcal{K} \otimes \mu_p \tag{3.19}$$

since $\delta_\mathcal{K}$ is invertible in $\mathcal{O}_\mathcal{K}/p\mathcal{O}_\mathcal{K}$. Let $\tilde{E} \subset \tilde{C} = \overline{Ig}\backslash Ig$ be *this* standard component.

Theorem 3.5 (i) *The operator Θ maps $H^0(S, \mathcal{L}^k)$ to $H^0(S, \mathcal{L}^{k+p+1})$.*

(ii) *The effect of Θ on Fourier–Jacobi expansions is a "Tate twist". More precisely, let*

$$\widetilde{FJ}(f) = \sum_{m=0}^{\infty} c_m(f) \tag{3.20}$$

be the canonical Fourier–Jacobi expansion of f along $\tilde{E}$ (thus $c_m(f) \in H^0(\tilde{E}, \mathcal{N}^m)$). Then

$$\widetilde{FJ}(\Theta(f)) = M^{-1} \sum_{m=0}^{\infty} mc_m(f). \tag{3.21}$$

Here M (equal to $N|D_{\mathcal{K}}|$ or $2^{-1}N|D_{\mathcal{K}}|$) is the width of the cusp.

(iii) *If $f \in H^0(S, \mathcal{L}^k)$ and $g \in H^0(S, \mathcal{L}^l)$, then*

$$\Theta(fg) = f\Theta(g) + \Theta(f)g. \tag{3.22}$$

(iv) *$\Theta(h_{\bar{\Sigma}}f) = h_{\bar{\Sigma}}\Theta(f)$ (equivalently, $\Theta(h_{\bar{\Sigma}}) = 0$).*

Corollary 3.6 *The operator Θ extends to a derivation of the graded ring of modular forms mod p, and for any f, $\Theta(f)$ is a cusp form.*

Parts (iii) and (iv) of the theorem are clear from the construction. The proof of (i), that $\Theta(f)$ is in fact *holomorphic* along S_{ss}, will be given in 3.4. We shall now study its effect on Fourier–Jacobi expansions, i.e. part (ii). That a factor like M^{-1} is necessary in (ii) becomes evident if we consider what happens to FJ expansions under level change. If N is replaced by $N' = NQ$, then the conormal bundle becomes the Q-th power of the conormal bundle of level N', i.e. $\mathcal{N} = \mathcal{N}'^Q$ (see Sect. 1.4.3). It follows that what was the m-th FJ coefficient at level N becomes the Qm-th coefficient at level N'. The operator Θ commutes with level change, but the factor M^{-1}, which changes to $(QM)^{-1}$, takes care of this.

3.3 The effect of Θ on FJ expansions

Let E be the standard cuspidal component of $\bar{S}$ (over the ring R_N). We have earlier trivialized the line bundle $\mathcal{L}$ along E in two seemingly different ways, that we must now compare. On the one hand, after reducing modulo $\mathfrak{P}$ (the prime of R_N above p fixed above) and pulling $\mathcal{L}$ back to the Igusa surface, we got a *canonical* nowhere vanishing section a trivializing $\mathcal{L}$ over $\overline{Ig}_{\mu}$, and in particular along any of the p^2-1 cuspidal components lying over E in $\overline{Ig}_{\mu}$. Using $\tilde{E}$ as a reference, there is a unique section of $\mathcal{L}$ along E which pulls back to $a|_{\tilde{E}}$. On the other hand, extending scalars from R_N to $\mathbb{C}$, shifting to the analytic category, restricting to the connected component $\bar{X}_{\Gamma}$ on which E lies, and then pulling back to a neighborhood of the cusp c_{∞} in the unit ball $\mathfrak{X}$, we have trivialized $\mathcal{L}|_E$ by means of the section $2\pi i d\zeta_3$, which we showed to be $\mathcal{K}_N$-rational.

Lemma 3.7 *The sections $a|_{\tilde{E}}$ and $2\pi i d\zeta_3$ "coincide" in the sense that they come from the same section in $H^0(E, \mathcal{L})$.*

Proof Let A be the universal semi-abelian variety over E. Its toric part is $\mathcal{O}_{\mathcal{K}} \otimes \mathbb{G}_m$, hence, taking $\bar{\Sigma}$-component of the cotangent space at the origin

$$\mathcal{L}|_{\tilde{E}} = \omega_{A/\tilde{E}}(\bar{\Sigma}) = (\delta_{\mathcal{K}}^{-1}\mathcal{O}_{\mathcal{K}} \otimes \omega_{\mathbb{G}_m})(\bar{\Sigma}) \tag{3.23}$$

admits the canonical section $e_{\bar{\Sigma}} \cdot (1 \otimes dT/T)$. Tracing back the definitions and using (1.69), this section becomes, under the base change $R_N \hookrightarrow \mathbb{C}$, just $2\pi i d\zeta_3$. On the other hand, when we reduce it modulo $\mathfrak{P}$ and use the Igusa level structure ε at the standard cusp, it pulls back to the section "with the same name" $e_{\bar{\Sigma}} \cdot (1 \otimes dT/T)$, because along $\tilde{E}$ (3.18) induces the identity on cotangent spaces. The lemma follows from the fact that, by definition, $\varepsilon^* a = e_{\bar{\Sigma}} \cdot (1 \otimes dT/T)$ too.

Lemma 3.8 *The sections* $a|_{\tilde{E}}^{p+1}$ *and* $2\pi i d\zeta_2 \otimes 2\pi i d\zeta_3 \mod \mathcal{P}_0 \otimes \mathcal{L}$ *"coincide" in the sense that they come from the same section in* $H^0(E, \mathcal{P}_\mu \otimes \mathcal{L})$*.*

Proof Let σ_2 (resp. σ_3) be the $\mathcal{K}_N$-rational section of $\mathcal{P}_\mu$ (resp. $\mathcal{L}$) along E, which over $\mathbb{C}$ becomes the section $2\pi i d\zeta_2$ (resp. $2\pi i d\zeta_3$). We have just seen that modulo $\mathfrak{P}$, when we identify $\tilde{E}$ with E (via the covering map $\tau : \overline{Ig} \to \bar{S}$), σ_3 reduces to a. To conclude, we must show that the map

$$V : \mathcal{P}/\mathcal{P}_0 = \mathcal{P}_\mu \simeq \mathcal{L}^{(p)} \tag{3.24}$$

carries σ_2 to $\sigma_3^{(p)}$. This will map, under $\mathcal{L}^{(p)} \simeq \mathcal{L}^p$, to a^p. Along E, the line bundles $\mathcal{P}_\mu$ and $\mathcal{L}$ are just the Σ- and $\bar{\Sigma}$-parts of the cotangent space at the origin of the torus $\mathcal{O}_\mathcal{K} \otimes \mathbb{G}_m$, and σ_2 and σ_3 are the sections

$$\sigma_2 = e_\Sigma \cdot (1 \otimes dT/T), \quad \sigma_3 = e_{\bar{\Sigma}} \cdot (1 \otimes dT/T). \tag{3.25}$$

Since in characteristic p, $V = Ver^* : \omega_{\mathbb{G}_m} \to \omega_{\mathbb{G}_m}^{(p)}$ maps dT/T to $(dT/T)^{(p)}$, for the $\mathcal{O}_\mathcal{K}$-torus, $V(\sigma_2) = \sigma_3^{(p)}$, and we are done.

To prove part (ii) of the main theorem, we argue as follows. Let $g = f/a^k$ be the function on $\overline{Ig}_\mu$ obtained by trivializing the line bundle $\mathcal{L}$. We have to study the FJ expansion along $\tilde{E}$ of $\psi(dg)/a^{p+1}$, where ψ is the map defined in (3.13). For that purpose, we may restrict to a formal neighborhood of $\tilde{E}$. This formal neighborhood is isomorphic, under the covering map $\tau : \overline{Ig}_\mu \to \bar{S}_\mu$, to the formal neighborhood $\widehat{S}$ of E in S. We may therefore regard dg as an element of $\Omega^1_{\widehat{S}}$. Now

$$\psi : \Omega^1_{\widehat{S}} \to \mathcal{P}_\mu \otimes \mathcal{L} \tag{3.26}$$

is a homomorphism of $\mathcal{O}_{\widehat{S}}$-modules defined over R_N so, having restricted to $\widehat{S}$, we may study the effect of ψ on FJ expansions by embedding $\widehat{S}_\mathbb{C}$ in a tubular neighborhood $\bar{S}(\varepsilon)$ of E and using complex analytic Fourier–Jacobi expansions. We are thus reduced to a complex-analytic computation, near the standard cusp at infinity.

Let

$$g(z, u) = \sum_{m=0}^{\infty} \theta_m(u) q^m \tag{3.27}$$

where $q = e^{2\pi i z/M}$ and θ_m is a theta function, so that $\theta_m(u)q^m$ is a section of $\mathcal{N}^m$ along E (now over $\mathbb{C}$). Then

$$\mathrm{d}g = 2\pi i M^{-1} \sum_{m=0}^{\infty} m\theta_m(u) q^m \mathrm{d}z + \sum_{m=0}^{\infty} \theta'_m(u) q^m \mathrm{d}u. \tag{3.28}$$

According to Corollary 1.30, $\psi(\mathrm{d}u) = 0$, and $\psi(\mathrm{d}z) = 2\pi i \mathrm{d}\zeta_2 \otimes \mathrm{d}\zeta_3$. It follows that

$$\psi(\mathrm{d}g) = M^{-1} \sum_{m=0}^{\infty} m\theta_m(u) q^m \cdot 2\pi i \mathrm{d}\zeta_2 \otimes 2\pi i \mathrm{d}\zeta_3. \tag{3.29}$$

Recalling that in characteristic p, $2\pi i \mathrm{d}\zeta_2 \otimes 2\pi i \mathrm{d}\zeta_3$ reduced to a^{p+1}, the proof of part (ii) of the theorem is now complete. For the convenience of the reader, we summarize the transitions between complex and p-adic maps in the following diagram:

$$\begin{array}{lllll}
/\kappa & \Omega^1_{\bar{S}_\mu/\kappa} & \overset{\mathrm{KS}(\Sigma)^{-1}}{\to} & \mathcal{P}\otimes\mathcal{L} & \\
& \cap & & \downarrow \quad \mathrm{mod}\ \mathcal{P}_0 & \\
/\kappa & \Omega^1_{\hat{S}/\kappa} & \overset{\psi}{\to} & \mathcal{P}_\mu\otimes\mathcal{L} & \overset{V\otimes 1}{\simeq}\ \mathcal{L}^{p+1} \\
& \uparrow \quad \mathrm{mod}\ \mathfrak{P} & & \uparrow & \\
/R_N & \Omega^1_{\hat{S}/R_N} & \overset{\psi}{\to} & \mathcal{P}_\mu\otimes\mathcal{L} & . \\
& \downarrow \otimes_{R_N}\mathbb{C} & & \downarrow & \\
/\mathbb{C} & \Omega^1_{\hat{S}/\mathbb{C}} & \overset{\psi}{\to} & \mathcal{P}_\mu\otimes\mathcal{L} & \\
& \cup & & \uparrow \quad \mathrm{mod}\ \mathcal{P}_0 & \\
/\mathbb{C} & \Omega^1_{\bar{S}(\varepsilon)/\mathbb{C}} & \overset{\mathrm{KS}(\Sigma)^{-1}_{an}}{\to} & \mathcal{P}\otimes\mathcal{L} &
\end{array} \tag{3.30}$$

We next turn to part (i).

3.4 A study of the theta operator along the supersingular locus

3.4.1 De Rham cohomology in characteristic p

We continue to consider the Picard surface S over κ and recall some facts about de Rham cohomology in characteristic p. Let $U = \mathrm{Spec}(R) \hookrightarrow S$ be a closed point s_0 ($R = \kappa = \mathcal{O}_{S,s_0}/\mathfrak{m}_{S,s_0}$), a nilpotent thickening of a closed point, or an affine open subset of S. We consider the restriction of the universal abelian scheme to R and denote it by A/R. Let $A^{(p)} = R \otimes_{\phi,R} A$ be its base change with respect to the map $\phi(x) = x^p$. Let

$$D = H^1_{dR}(A/R), \tag{3.31}$$

a locally free R-module of rank 6. The de Rham cohomology of $A^{(p)}$ is

$$D^{(p)} = R \otimes_{\phi,R} D. \tag{3.32}$$

The R-linear Frobenius and Verschiebung morphisms $Frob : A \to A^{(p)}$, $Ver : A^{(p)} \to A$ induce (by pull-back) linear maps

$$F : D^{(p)} \to D,\ \ V : D \to D^{(p)}. \tag{3.33}$$

Both F and V are everywhere of rank 3, which implies that their kernel and image are locally free direct summands. Moreover, $\mathrm{Im}F = \ker V$ and $\mathrm{Im}V = \ker F = \omega_{A^{(p)}/R}$. The maps F and V preserve the types $\Sigma, \bar{\Sigma}$, but note that $D^{(p)}(\Sigma) = D(\bar{\Sigma})^{(p)}$ etc.

The principal polarization on A induces one on $A^{(p)}$, and these polarizations induce symplectic forms

$$\langle , \rangle : D \times D \to R,\ \ \langle , \rangle^{(p)} : D^{(p)} \times D^{(p)} \to R \tag{3.34}$$

where the second form is just the base change of the first. For $x \in D^{(p)}, y \in D$, we have

$$\langle Fx, y\rangle = \langle x, Vy\rangle^{(p)}. \tag{3.35}$$

In addition, for $a \in \mathcal{O}_\mathcal{K}$

$$\langle \iota(a)x, y\rangle = \langle x, \iota(\bar{a})y\rangle. \tag{3.36}$$

As $VF = FV = 0$, the first relation implies that $\mathrm{Im}F$ and $\mathrm{Im}V$ are isotropic subspaces. So is $\omega_{A/R}$.

The Gauss–Manin connection is an integrable connection

$$\nabla : D \to \Omega^1_R \otimes D. \tag{3.37}$$

It is a priori defined (e.g. in [24]) when R is smooth over κ, but we can define it by base change also when R is a nilpotent thickening of a point of S (see [25], where R is a local Artinian ring).

We shall need to deal only with the first infinitesimal neighborhood of a point, $R = \mathcal{O}_{S,s_0}/\mathfrak{m}^2_{S,s_0}$. In this case, D has a basis of horizontal sections. Indeed, $R = \kappa[u,v]/(u^2, uv, v^2)$ where u and v are local parameters at s_0, and

$$\Omega^1_R = (R\mathrm{d}u + R\mathrm{d}v)/\left\langle u\mathrm{d}u,\ v\mathrm{d}v,\ u\mathrm{d}v + v\mathrm{d}u\right\rangle$$

(p is odd). If $x \in D$ and

$$\nabla x = \mathrm{d}u \otimes x_1 + \mathrm{d}v \otimes x_2 \tag{3.38}$$

then $\tilde{x} = x - ux_1 - vx_2$ is horizontal, so the horizontal sections span D over R by Nakayama's lemma. It follows that if $D_0 = D^\nabla$ is the space of horizontal sections,

$$R \otimes_\kappa D_0 = D, \tag{3.39}$$

$\nabla = d \otimes 1$ and we can identify $D_0 = H^1_{dR}(A_{s_0}/\kappa)$, i.e. every de Rham class at s_0 has a *unique* extension to a horizontal section $x \in H^1_{dR}(A/R)$.

There is a similar connection on $D^{(p)}$. The isogenies *Frob* and *Ver*, like any isogeny, take horizontal sections with respect to the Gauss–Manin connection to horizontal sections, e.g. if $x \in D$ and $\nabla x = 0$, then $Vx \in D^{(p)}$ satisfies $\nabla(Vx) = 0$.

The pairing $\langle , \rangle$ is horizontal for ∇, i.e.

$$\mathrm{d}\left\langle x, y\right\rangle = \left\langle \nabla x, y\right\rangle + \left\langle x, \nabla y\right\rangle. \tag{3.40}$$

Remark 3.1 In the theory of Dieudonné modules, one works over a perfect base. It is then customary to identify D with $D^{(p)}$ via $x \leftrightarrow 1 \otimes x$. This identification is only σ-linear where $\sigma = \phi$, now viewed as an *automorphism* of R. The operator F becomes σ-linear, V becomes σ^{-1}-linear, and (3.35) reads $\left\langle Fx, y\right\rangle = \left\langle x, Vy\right\rangle^\sigma$. With this convention, F and V switch types, rather than preserve them.

3.4.2 The Dieudonné module at a gss point

Assume from now on that $s_0 \in Z' = S_{\mathrm{gss}}$ is a closed point of the general supersingular locus. We write D_0 for $H^1_{dR}(A_{s_0}/\kappa)$.

Lemma 3.9 *There exists a basis $e_1, e_2, f_3, f_1, f_2, e_3$ of D_0 with the following properties. Denote by $e_1^{(p)} = 1 \otimes e_1 \in D_0^{(p)}$ etc.*

- (i) *$\mathcal{O}_\mathcal{K}$ acts on the e_i via Σ and on the f_i via $\bar{\Sigma}$ (hence it acts on the $e_i^{(p)}$ via $\bar{\Sigma}$ and on the $f_i^{(p)}$ via Σ).*
- (ii) *The symplectic pairing satisfies*

$$\left\langle e_i, f_j\right\rangle = -\left\langle f_j, e_i\right\rangle = \delta_{ij},\quad \left\langle e_i, e_j\right\rangle = \left\langle f_i, f_j\right\rangle = 0. \tag{3.41}$$

- (iii) *The vectors e_1, e_2, f_3 form a basis for the cotangent space $\omega_{A_0/\kappa}$. Hence e_1 and e_2 span $\mathcal{P}$ and f_3 spans $\mathcal{L}$.*
- (iv) $\ker(V)$ *is spanned by e_1, f_2, e_3. Hence $\mathcal{P}_0 = \mathcal{P} \cap \ker(V)$ is spanned by e_1.*
- (v) $Ve_2 = f_3^{(p)},\ Vf_3 = e_1^{(p)},\ Vf_1 = e_2^{(p)}$.
- (vi) $Ff_1^{(p)} = -e_3,\ Ff_2^{(p)} = -e_1,\ Fe_3^{(p)} = -f_2$.

Proof Up to a slight change of notation, this is the unitary Dieudonné module which Bültel and Wedhorn call a "braid of length 3" and denote by $\bar{B}(3)$, cf [4] (3.2). The classification in loc. cit. Proposition 3.6 shows that the Dieudonné module of a μ-ordinary abelian variety is isomorphic to $\bar{B}(2) \oplus \bar{S}$, that of a gss abelian variety is isomorphic to $\bar{B}(3)$ and in the superspecial case we get $\bar{B}(1) \oplus \bar{S}^2$.

3.4.3 Infinitesimal deformations

Let $\mathcal{O}_{S,s_0}$ be the local ring of S at s_0, $\mathfrak{m}$ its maximal ideal, and $R = \mathcal{O}_{S,s_0}/\mathfrak{m}^2$. This R is a truncated polynomial ring in two variables, isomorphic to $\kappa[u, v]/(u^2, uv, v^2)$.

As remarked above, the de Rham cohomology $D = H^1_{dR}(A/R)$ has a basis of horizontal sections and we may identify D^∇ with D_0 and D with $R \otimes_\kappa D_0$.

Grothendieck tells us that A/R is completely determined by A_0 and by the Hodge filtration $\omega_{A/R} \subset D = R \otimes_\kappa D_0$. Since A *is* the universal infinitesimal deformation of A_0, we may choose the coordinates u and v so that

$$\mathcal{P} = \mathrm{Span}_R\{e_1 + ue_3, e_2 + ve_3\}. \tag{3.42}$$

The fact that $\omega_{A/R}$ is isotropic implies then that

$$\mathcal{L} = \mathrm{Span}_R\{f_3 - uf_1 - vf_2\}. \tag{3.43}$$

Consider the abelian scheme $A^{(p)}$. It is *not* the universal deformation of $A_0^{(p)}$ over R. In fact, the map $\phi : R \to R$ factors as

$$R \xrightarrow{\pi} \kappa \xrightarrow{\phi} \kappa \xrightarrow{i} R, \tag{3.44}$$

and therefore $A^{(p)}$, *unlike* A, is constant: $A^{(p)} = \mathrm{Spec}(R) \times_{\mathrm{Spec}(\kappa)} A_0^{(p)}$. As with D, $D^{(p)} = R \otimes_\kappa D_0^{(p)}$, $\nabla = d \otimes 1$, but this time the basis of horizontal sections can be obtained also from the trivialization of $A^{(p)}$, and $\omega_{A^{(p)}/R} = \mathrm{Span}_R\{e_1^{(p)}, e_2^{(p)}, f_3^{(p)}\}$.

Since V and F preserve horizontality, e_1, f_2, e_3 span $\ker(V)$ over R in D, and the relations in (v) and (vi) of Lemma 3.9 continue to hold. Indeed, the matrix of V in the basis at s_0 prescribed by that lemma, continues to represent V over $\mathrm{Spec}(R)$ by "horizontal continuation". The matrix of F is then derived from the relation (3.35).

The Hodge filtration nevertheless varies, so we conclude that

$$\mathcal{P}_0 = \mathcal{P} \cap \ker(V) = \mathrm{Span}_R\{e_1 + ue_3\}. \tag{3.45}$$

The condition $V(\mathcal{L}) = \mathcal{P}_0^{(p)}$, which is the "equation" of the closed subscheme $Z' \cap \mathrm{Spec}(R)$ (see Theorem 2.9) means

$$V(f_3 - uf_1 - vf_2) = e_1^{(p)} - ue_2^{(p)} \in R \cdot e_1^{(p)} \tag{3.46}$$

and this holds if and only if $u = 0$. We have proved the following lemma.

Lemma 3.10 *Let $s_0 \in S_{\mathrm{gss}}$ and the notation be as above. Then the closed subscheme $S_{\mathrm{gss}} \cap \mathrm{Spec}(R)$ is given by the equation $u = 0$.*

3.4.4 The Kodaira–Spencer isomorphism along the general supersingular locus

We keep the assumptions of the previous subsections and compute what the Gauss–Manin connection does to $\mathcal{P}_0$. A typical element of $\mathcal{P}_0$ is $g(e_1 + ue_3)$ for some $g \in R$. Then

$$\nabla(g(e_1 + ue_3)) = \mathrm{d}g \otimes (e_1 + ue_3) + g\mathrm{d}u \otimes e_3. \tag{3.47}$$

Note that when we divide by $\omega_{A/R}$ and project $H^1_{dR}(A/R)$ to $H^1(A,\mathcal{O})$, e_1+ue_3 dies, and the image $\overline{e_3}$ of e_3 becomes a basis for the line bundle that we called $\mathcal{L}^\vee(\rho)=H^1(A,\mathcal{O})(\Sigma)$. Recall the definition of ψ given in (3.13), but note that this definition only makes sense over $\mathrm{Spec}(\mathcal{O}_{S,s_0})$ or its completion, where $\mathrm{KS}(\Sigma)$ is an isomorphism, and can be inverted.

Proposition 3.11 *Let $s_0\in Z'=S_{\mathrm{gss}}$. Choose local parameters u and v at s_0 so that in $\mathcal{O}_{S,s_0}$ the local equation of Z' becomes $u=0$. Then at s_0, $\psi(\mathrm{d}u)$ has a zero along Z'.*

Proof Let $i:Z'\hookrightarrow S$ be the locally closed embedding. We must show that in a suitable Zariski neighborhood of s_0, where $u=0$ is the local equation of Z', $i^*\psi(\mathrm{d}u)=0$. It is enough to show that the image of $\psi(\mathrm{d}u)$ in the *fiber* at every point s of Z' near s_0, vanishes. All points being alike, it is enough to do it at s_0. In other words, we denote by ψ_0 the map

$$\psi_0:\Omega^1_{S,s_0}\to\mathcal{P}_\mu\otimes\mathcal{L}|_{s_0}\simeq\mathcal{L}^{p+1}|_{s_0}. \tag{3.48}$$

and show that $\psi_0(\mathrm{d}u)=0$. We may now work over $\mathrm{Spec}(R)$, where $R=\mathcal{O}_{S,s_0}/\mathfrak{m}^2$. It is enough to show that in the diagram

$$\begin{array}{lll}\mathcal{P}_R\otimes\mathcal{L}_R & \overset{\mathrm{KS}(\Sigma)}{\to} & \Omega^1_R\\ \downarrow & & \downarrow\\ \mathcal{P}_{s_0}\otimes\mathcal{L}_{s_0}\simeq & & \Omega^1_{S,s_0}\end{array} \tag{3.49}$$

$\mathrm{KS}(\Sigma)$ maps the line sub-bundle $\mathcal{P}_{0,R}\otimes\mathcal{L}_R$ onto $R\mathrm{d}u$. Once we have passed to the infinitesimal neighborhood $\mathrm{Spec}(R)$, we can replace the local parameters u,v by any two formal parameters for which $u=0$ defines $Z'\cap\mathrm{Spec}(R)$. We may therefore assume, in view of Lemma 3.10, that u and v have been chosen as in Sect. 3.4.3. But then (3.47) shows that the restriction of $\mathrm{KS}(\Sigma)$ to Z', i.e. the homomorphism $i^*\mathrm{KS}(\Sigma)$, maps $i^*\mathcal{P}_0$ onto $i^*R\cdot\mathrm{d}u\otimes\overline{e_3}$. This concludes the proof.

3.4.5 A computation of poles along the supersingular locus

We are now ready to prove the following.

Proposition 3.12 *Let $k\geq 0$, and let $f\in H^0(S,\mathcal{L}^k)$ be a modular form of weight k in characteristic p. Then $\Theta(f)\in H^0(S,\mathcal{L}^{k+p+1})$.*

Proof A priori, the definition that we have given for $\Theta(f)$ produces a meromorphic section of $\mathcal{L}^{k+p+1}$ which is holomorphic on the μ-ordinary part S_μ but may have a pole along $Z=S_{\mathrm{ss}}$. Since S is a non-singular surface, it is enough to show that $\Theta(f)$ does not have a pole along $Z'=S_{\mathrm{gss}}$, the non-singular part of the divisor Z. Consider the degree p^2-1 covering $\tau:Ig\to S$, which is finite, étale over S_μ and totally ramified along Z. Let $s_0\in Z'$ and let $\tilde{s}_0\in Ig$ be the closed point above it. Let u,v be formal parameters at s_0 for which Z' is given by $u=0$, as in Theorem 2.15. As explained there, we may choose formal parameters w,v at $\tilde{s}_0$ where $w^{p^2-1}=u$ (and v is the same function v pulled back from S to Ig). It follows that in Ω^1_{Ig} we have

$$\mathrm{d}u=-w^{p^2-2}\mathrm{d}w. \tag{3.50}$$

We now follow the steps of our construction. Dividing f by a^k, we get a function $g=f/a^k$ on Ig with a pole of order at most k along $\tilde{Z}$, the supersingular divisor on Ig, whose local equation is $w=0$. In $\widehat{\mathcal{O}}_{Ig,\tilde{s}_0}$ we may write

$$g = \sum_{l=-k}^{\infty} g_l(v)w^l. \tag{3.51}$$

Then

$$\begin{aligned} \mathrm{d}g &= \sum_{l=-k}^{\infty} lg_l(v)w^{l-1}\mathrm{d}w + \sum_{l=-k}^{\infty} w^l g_l'(v)\mathrm{d}v \\ &= -\sum_{l=-k}^{\infty} lg_l w^{l-(p^2-1)}\mathrm{d}u + \sum_{l=-k}^{\infty} w^l g_l'(v)\mathrm{d}v. \end{aligned} \tag{3.52}$$

Applying the map ψ (extended $\mathcal{O}_{Ig}$-linearly from S to Ig), and noting that $\psi(\mathrm{d}u)$ has a zero along Z', hence a zero of order p^2-1 along $\tilde{Z}'$, we conclude that $\psi(\mathrm{d}g)$ has a pole of order k (at most) along $\tilde{Z}'$. Finally, $\Theta(f) = a^k \cdot \psi(\mathrm{d}g)$ becomes holomorphic along $\tilde{Z}'$, and also descends to S. It is therefore a holomorphic section of $\mathcal{P}_\mu \otimes \mathcal{L}^{k+1} \simeq \mathcal{L}^{k+p+1}$.

It is amusing to compare the reasons for the increase by $p+1$ in the weight of $\Theta(f)$ for modular curves and for Picard modular surfaces. In the case of modular curves the Kodaira–Spencer isomorphism is responsible for a shift by 2 in the weight, but the section acquires simple poles at the supersingular points. One has to multiply it by the Hasse invariant, which has weight $p-1$, to make the section holomorphic and hence a total increase by $p+1 = 2+(p-1)$ in the weight. In our case, the map ψ is responsible for a shift by $p+1$ (the p coming from $\mathcal{P}_\mu \simeq \mathcal{L}^p$), but the section turns out to be holomorphic along the supersingular locus. See Section 4.2.

4 Further results on Θ

4.1 Relation to the filtration and theta cycles

In part (ii) of Theorem 3.5, we have described the way Θ acts on Fourier–Jacobi expansions at the standard cusp. A similar formula holds at all the other cusps. We deduce from it that modular forms in the image of Θ have vanishing FJ coefficients in degrees divisible by p. Moreover, for such a form $f \in \mathrm{Im}(\Theta)$, $\Theta^{p-1}(f)$ and f have the same FJ expansions, and hence the same filtration. Note also that if $r(f_1) = r(f_2)$, then $r(\Theta(f_1)) = r(\Theta(f_2))$. We may therefore define unambiguously

$$\Theta(r(f)) = r(\Theta(f)). \tag{4.1}$$

As we clearly have

$$\omega(\Theta(f)) = \omega(f) + p + 1 - a\left(p^2-1\right) \tag{4.2}$$

for some $a \geq 0$ we deduce the following result.

Proposition 4.1 *Let $f \in M_k(N,\kappa)$ be a modular form modulo p, and assume that $r(f) \in Im(\Theta)$. Then*

$$r(f) = r\left(\Theta^{p-1}(f)\right). \tag{4.3}$$

There exists a unique index $0 \leq i \leq p-2$ such that

$$\omega(\Theta^{i+1}(f)) = \omega\left(\Theta^i(f)\right) + p + 1 - \left(p^2-1\right). \tag{4.4}$$

For any other i in this range

$$\omega(\Theta^{i+1}(f)) = \omega\left(\Theta^i(f)\right) + p + 1. \tag{4.5}$$

This is reminiscent of the "theta cycles" for classical (i.e. elliptic) modular forms modulo p, see [19,21,34]. Recall that if f is a $\mod p$ modular form of weight k on $\Gamma_0(N)$ with q-expansion $\sum a_n q^n$ ($a_n \in \bar{\mathbb{F}}_p$), then $\theta(f)$ is a $\mod p$ modular form of weight $k+p+1$ with q-expansion $\sum na_n q^n$ (Katz denotes $\theta(f)$ by $A\theta(f)$). One has $\omega(\theta(f)) < \omega(f)+p+1$ *if and only if* $\omega(f) \equiv 0 \mod p$. In such a case, we say that the filtration "drops" and we have

$$\omega(\theta(f)) = \omega(f) + p + 1 - a(p-1) \tag{4.6}$$

for some $a > 0$. As a corollary, $\omega(f)$ can never equal $1 \mod p$ for an $f \in \mathrm{Im}(\theta)$. Assume now that $f \in \mathrm{Im}(\theta)$ is a "low point" in its "theta cycle", namely, $\omega(f)$ is minimal among all $\omega(\theta^i(f))$. Then $\omega(\theta^{i+1}(f)) < \omega(\theta^i(f)) + p + 1$ for *one* or *two* values of $i \in [0, p-2]$, which are completely determined by $\omega(f) \mod p$ [19].

This is not true anymore for Picard modular forms. Not only is the drop in the theta cycle unique, but the question of when exactly it occurs is mysterious and deserves further study. We make the following elementary observation showing that whether a drop in the filtration occurs in passing from f to $\Theta(f)$ can *not* be determined by $\omega(f)$ modulo p alone. Let f and k be as in Proposition 4.1.

(1) If $k \leq p^2 - 1$, then $\omega(f) = k$.
(2) If $k < p+1$, then $\omega(\Theta^i(f)) = k + i(p+1)$ for $0 \leq i \leq p-2$, so the drop occurs at the last step of the theta cycle, i.e. at weight $k + (p-2)(p+1)$, which is congruent to $k-2$ modulo p.
(3) If $k < p+1$ but $r(f) \notin \mathrm{Im}(\Theta)$, then starting with $\Theta(f)$ instead of f, one sees that the drop in the theta cycle of $\Theta(f)$ occurs either in passing from $\Theta^{p-2}(f)$ to $\Theta^{p-1}(f)$, or in passing from $\Theta^{p-1}(f)$ to $\Theta^p(f)$.

4.2 Compatibility between theta operators for elliptic and Picard modular forms

4.2.1 The theta operator for elliptic modular forms

The theta operator for elliptic modular forms modulo p was introduced by Serre and Swinnerton-Dyer in terms of q-expansions, but its geometric construction was given by Katz [20,21]. Katz relied on a canonical splitting of the Hodge filtration over the ordinary locus, but Gross gave in [13], Proposition 5.8, the construction after which we modeled our Θ.

Let us quickly repeat Gross' construction as outlined in the introduction. Let X be the open modular curve $X(N)$ over $\bar{\mathbb{F}}_p$ ($N \geq 3$, $p \nmid N$) and I_{ord} the Igusa curve of level p lying over $X_{\mathrm{ord}} = X \backslash X_{\mathrm{ss}}$, the ordinary part of X. Let $\bar{X}$ and $\bar{I}_{\mathrm{ord}}$ be the curves obtained by adjoining the cusps to X and I_{ord}, respectively. Let $\mathcal{L} = \omega_{E/X}$ be the cotangent bundle of the universal elliptic curve, extended over the cusps as usual. Classical modular forms of weight k and level N are sections of $\mathcal{L}^k$ over $\bar{X}$. Let a be the tautological nowhere vanishing section of $\mathcal{L}$ over $\bar{I}_{\mathrm{ord}}$. Given a modular form f of weight k, we consider $r(f) = \tau^* f / a^k$ where $\tau : \bar{I}_{\mathrm{ord}} \to \bar{X}$ is the covering map, and apply the inverse of the Kodaira–Spencer isomorphism $\mathrm{KS} : \mathcal{L}^2 \to \Omega^1_{\bar{I}_{\mathrm{ord}}}$ to get a section $\mathrm{KS}^{-1}(\mathrm{d}r(f))$ of $\mathcal{L}^2$ over $\bar{I}_{\mathrm{ord}}$. When multiplied by a^k it descends to $\bar{X}_{\mathrm{ord}}$, and when this is multiplied further by $h = a^{p-1}$, the Hasse invariant for elliptic modular forms, it extends holomorphically over X_{ss} to an element

$$\theta(f) = a^{k+p-1}\mathrm{KS}^{-1}(\mathrm{d}r(f)) \in H^0\left(\bar{X}, \mathcal{L}^{k+p+1}\right). \tag{4.7}$$

4.2.2 An embedding of a modular curve in $\bar{S}$

To illustrate our idea, and to simplify the computations, we assume that $N = 1$ and $d_{\mathcal{K}} \equiv 1 \mod 4$, so that $D = D_{\mathcal{K}} = d_{\mathcal{K}}$. This conflicts of course with our running hypothesis $N \geq 3$, but for the current section does not matter much. We shall treat only one special embedding of the modular curve $\bar{X} = X_0(D)$ into $\bar{S}$ (there are many more).

Embed $SL_2(\mathbb{R}) = SU(1,1)$ in G'_∞ via

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} a & & b \\ & 1 & \\ c & & d \end{pmatrix}. \tag{4.8}$$

This embedding induces an embedding of symmetric spaces $\mathfrak{H} \hookrightarrow \mathfrak{X}$, $z \mapsto {}^t(z,0)$. One can easily compute that the intersection of Γ, the stabilizer of the lattice L_0 in G'_∞, with $SL_2(\mathbb{R})$, is the subgroup of $SL_2(\mathbb{Z})$ given by

$$\Gamma^0(D) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : D|b \right\}. \tag{4.9}$$

Let $E_0 = \mathbb{C}/\mathcal{O}_{\mathcal{K}}$, endowed with the canonical principal polarization and CM type Σ. For $z \in \mathfrak{H}$, let $\Lambda_z = \mathbb{Z} + \mathbb{Z}z$ and $E_z = \mathbb{C}/\Lambda_z$. Let M_z be the cyclic subgroup of order D of E_z generated by $D^{-1}z \mod \Lambda_z$. Using the model (1.27) of the abelian variety A_z associated with the point ${}^t(z,0) \in \mathfrak{X}$, we compute that

$$A_z \simeq E_0 \times (\mathcal{O}_{\mathcal{K}} \otimes E_z)/(\delta_{\mathcal{K}} \otimes M_z) \tag{4.10}$$

with the obvious $\mathcal{O}_{\mathcal{K}}$-structure. The group $\delta_{\mathcal{K}} \otimes M_z$ is a cyclic subgroup of $\mathcal{O}_{\mathcal{K}} \otimes E_z$ of order D, generated by $\delta_{\mathcal{K}}^{-1} \otimes z \mod \mathcal{O}_{\mathcal{K}} \otimes \Lambda_z$. The principal polarization on A_z provided by the complex uniformization is the product of the canonical polarization of E_0 and the principal polarization of $\mathcal{O}_{\mathcal{K}} \otimes E_z/\delta_{\mathcal{K}} \otimes M_z$ obtained by descending the polarization

$$\lambda_{can} : \mathcal{O}_{\mathcal{K}} \otimes E_z \to \delta_{\mathcal{K}}^{-1} \otimes E_z = (\mathcal{O}_{\mathcal{K}} \otimes E_z)^t \tag{4.11}$$

of degree D^2, modulo the maximal isotropic subgroup $\delta_{\mathcal{K}} \otimes M_z$ of $\ker(\lambda_{can})$.

It is now clear that over any R_0-algebra R we have the same moduli-theoretic construction, sending a pair (E, M) where M is a cyclic subgroup of degree D to $A(E, M)$, with $\mathcal{O}_{\mathcal{K}}$ structure and polarization given by the same formulae. This gives a modular embedding $j : X \to S$ which is generically injective. To make this precise at the level of schemes (rather than stacks), one would have to add a level N structure and replace the base ring R_0 by R_N.

4.2.3 Comparison of the two theta operators

From now on, we work over $\bar{\mathbb{F}}_p$. The modular interpretation of the embedding $j : \bar{X} \to \bar{S}$ allows us to complete it to a diagram

$$\begin{array}{ccc} \bar{I}_{\mathrm{ord}} & \xrightarrow{j} & \overline{Ig}_\mu \\ \tau \downarrow & & \downarrow \tau\,. \\ \bar{X}_{\mathrm{ord}} & \xrightarrow{j} & \bar{S}_\mu \end{array} \tag{4.12}$$

Note that $j(X_{\mathrm{ss}}) \subset S_{\mathrm{ssp}}$, i.e. the embedded modular curve cuts the supersingular locus at superspecial points.

Lemma 4.2 *The pull-back $j^*\omega_{\mathcal{A}/S}$ decomposes as a product $\omega_{E_0} \times (\mathcal{O}_\mathcal{K} \otimes \omega_{E/X})$. Under this isomorphism,*

$$\begin{aligned} j^*\mathcal{L} &= (\mathcal{O}_\mathcal{K} \otimes \omega_{E/X})(\bar{\Sigma}) \\ j^*\mathcal{P}_0 &= \omega_{E_0} \\ j^*\mathcal{P}_\mu &\simeq (\mathcal{O}_\mathcal{K} \otimes \omega_{E/X})(\Sigma). \end{aligned} \tag{4.13}$$

The line bundle $j^*\mathcal{P}_0$ is constant, and $\mathcal{P}_\mu$, originally a quotient bundle of $\mathcal{P}$, becomes a direct summand when restricted to $\bar{X}$.

Proof This is straightforward from the construction of j, and the fact that E_0 is supersingular, while E is ordinary over $\bar{X}_{\mathrm{ord}}$. Note that $\mathcal{O}_\mathcal{K} \otimes E/\delta_\mathcal{K} \otimes M$ and $\mathcal{O}_\mathcal{K} \otimes E$ have the same cotangent space.

Proposition 4.3 *Identify $j^*\mathcal{L}$ with $\omega_{E/X}$ ($\mathcal{O}_\mathcal{K}$ acting via $\bar{\Sigma}$). Then for $f \in H^0(\bar{S}, \mathcal{L}^k) = M_k(N, \bar{\mathbb{F}}_p)$*

$$\theta(j^*(f)) = j^*(\Theta(f)). \tag{4.14}$$

Proof We abbreviate I_{ord} by I and Ig_μ by Ig. The pull-back via j of the tautological section a of $\mathcal{L}$ over Ig is the tautological section a of $j^*\mathcal{L} = \omega_{E/X}$. We therefore have

$$j^*(\mathrm{d}r(f)) = \mathrm{d}r(j^*(f)) \tag{4.15}$$

($r(f) = \tau^* f / a^k$ is the function on Ig denoted earlier also by g). It remains to check the commutativity of the following diagram

$$\begin{array}{ccccc} \Omega^1_{Ig} & \overset{\mathrm{KS}(\Sigma)^{-1}}{\rightarrow} & \mathcal{P} \otimes \mathcal{L} & \overset{V\otimes 1}{\rightarrow} & \mathcal{L}^{p+1} \\ \downarrow j_0^* & & & & \downarrow j^* \\ \Omega^1_I & \overset{\mathrm{KS}^{-1}}{\rightarrow} & j^*\mathcal{L}^2 & \overset{\times h}{\rightarrow} & j^*\mathcal{L}^{p+1} \end{array}. \tag{4.16}$$

Here j_0^* is the map $j^*\Omega^1_{Ig} \to \Omega^1_I$ on differentials whose kernel is the conormal bundle of I in Ig. For that we have to compare the Kodaira–Spencer maps on S and on X. As we have seen in the lemma, $\mathcal{P}/\mathcal{P}_0 = \mathcal{P}_\mu$ pulls back under j to $\mathcal{L}(\rho)$ (the line bundle $\mathcal{L}$ with the $\mathcal{O}_\mathcal{K}$ action conjugated). But, $\mathrm{KS}(\Sigma)(\mathcal{P}_0 \otimes \mathcal{L})$ maps under j^* to the conormal bundle, so we obtain a commutative diagram

$$\begin{array}{ccc} \Omega^1_{Ig} & \overset{\mathrm{KS}(\Sigma)}{\leftarrow} & \mathcal{P} \otimes \mathcal{L} \\ \downarrow j_0^* & & \downarrow \quad \bmod \mathcal{P}_0 \\ \Omega^1_I & \overset{\mathrm{KS}}{\leftarrow} & j^*\mathcal{L}(\rho) \otimes j^*\mathcal{L} \end{array}. \tag{4.17}$$

The commutativity of the diagram

$$\begin{array}{ccc} \mathcal{P}_\mu & \overset{V}{\rightarrow} & \mathcal{L}^{(p)} \\ \downarrow & & \downarrow \\ j^*\mathcal{L}(\rho) & \overset{\times h}{\rightarrow} & j^*\mathcal{L}^{(p)} \end{array} \tag{4.18}$$

follows from the definition of the Hasse invariant h on X. Identifying $\mathcal{L}^{(p)}$ with $\mathcal{L}^p$ as usual and tensoring the last diagram with $\mathcal{L}$ provides the last piece of the puzzle.

Remark 4.1 The proposition follows, of course, also from the effect of θ and Θ on q-expansions, once we compare FJ expansions on $\bar{S}$ to q-expansions on the embedded $\bar{X}$. The geometric proof given here has the advantage that it explains the precise way in which $V_\mathcal{P} \otimes 1$ replaces "multiplication by h''.

5 The Igusa tower and *p*-adic modular forms

We shall be very brief, since from now on the development follows closely the classical case of p-adic modular forms on $GL(2)$, with minor modifications. A general reference for this section is Hida's book [16], although, strictly speaking, our case (p inert) is excluded there.

5.1 Geometry modulo p^m

5.1.1 The Picard surface modulo p^m

Let $m \geq 1$, and write $R_m = R_0/p^m R_0 = \mathcal{O}_\mathcal{K}/p^m\mathcal{O}_\mathcal{K}$. Let

$$S^{(m)} = S \times_{\mathrm{Spec}(R_0)} \mathrm{Spec}(R_m) \tag{5.1}$$

so that $S^{(1)} = S_{\kappa_0}$ is the special fiber, and use a similar notation for the complete surface $\bar{S}^{(m)}$. Write $S_\mu^{(m)}$ (resp. $\bar{S}_\mu^{(m)}$) for the Zariski open subset of points whose image in $\bar{S}^{(1)}$ lies in $S_\mu^{(1)}$ (resp. in $\bar{S}_\mu^{(1)}$).

The generic fiber (in the sense of Raynaud) of the formal scheme

$$\varinjlim \bar{S}_\mu^{(m)} \tag{5.2}$$

is a rigid analytic space which we shall denote by $\bar{S}_\mu^{rig}$. We shall refer to its complement in $\bar{S}^{rig}$ (the rigid analytic space associated with $\bar{S}$) as the *supersingular tube.* Its $\mathbb{C}_p$-points are the points of $\bar{S}(\mathbb{C}_p)$ whose reduction modulo p lies in $S_{ss}(\bar{\mathbb{F}}_p)$.

5.1.2 p-Adic modular forms of integral weight k

The vector bundles $\mathcal{P}$ and $\mathcal{L}$ induce vector bundles on $\bar{S}^{(m)}$ and $\bar{S}_\mu^{rig}$ which we shall denote by the same symbols (the latter in the rigid analytic category). Let $k \in \mathbb{Z}$ (k may be negative). Let R be a topological $\mathcal{K}_p$-algebra. We define a *p-adic modular form of weight k and tame level N over R* to be an element f of

$$M_k^p(N;R) := H^0\left(\bar{S}_\mu^{rig}\widehat{\otimes}_{\mathcal{K}_p}R, \mathcal{L}^k\right). \tag{5.3}$$

Note that $M_k^p(N;R) = R\widehat{\otimes}_{\mathcal{K}_p}M_k^p(N;\mathcal{K}_p)$. A p-adic modular form f is said to be *overconvergent* if there exist finitely many $\mathcal{K}_p$-affinoids X_i contained in the supersingular tube and a section of $\mathcal{L}^k$ over $(\bar{S}^{rig}\setminus \bigcup X_i)\widehat{\otimes}_{\mathcal{K}_p}R$ which restricts to f. We denote the subspace of overconvergent modular forms by $M_k^{oc}(N;R)$.

Note that if R is not of topologically finite type over $\mathcal{K}_p$ our definition of "overconvergent" is a priori stronger than asking f to extend to a strict neighborhood of $\bar{S}_\mu^{rig}\widehat{\otimes}_{\mathcal{K}_p}R$ in $\bar{S}^{rig}\widehat{\otimes}_{\mathcal{K}_p}R$.

The space $M_k^p(N;\mathcal{K}_p)$ is a p-adic Banach space whose unit ball is given by

$$M_k^p(N;\mathcal{O}_p) = \varprojlim H^0\left(\bar{S}_\mu^{(m)}, \mathcal{L}^k\right). \tag{5.4}$$

5.1.3 q-Expansion principle

Whether we are dealing with an $f \in H^0(\bar{S}_\mu^{(m)}, \mathcal{L}^k)$ or an $f \in M_k^p(N;\mathcal{K}_p)$ the same procedure as in Sect. 1.10.2 allows us to associate with f a Fourier–Jacobi expansion $FJ(f)$ (1.125). Recall, however, that $FJ(f)$ depends on the section $s \in H^0(C,\mathcal{L})$ used to trivialize $\mathcal{L}|_C$. Note that if $f \in M_k^p(N;\mathcal{K}_p)$, the coefficients of $FJ(f)$ are theta functions with *bounded denominators*, since a suitable $\mathcal{K}_p$-multiple of f lies in $M_k^p(N;\mathcal{O}_p)$.

As with classical modular forms, we have the q-expansion principle, stemming from the fact that C meets every component of $\bar{S}_\mu^{rig}$.

Lemma 5.1 *If $FJ(f) = 0$, then $f = 0$.*

Corollary 5.2 *If $f \in M_k^p(N;\mathcal{O}_p)$ and $FJ(f)$ is divisible by p (in the sense that every $c_j(f) \in H^0(C,\mathcal{N}^j)$ is divisible by p with respect to the integral structure on $\bar{S}$), then $f \in pM_k^p(N;\mathcal{O}_p)$.*

5.2 The Igusa scheme of level p^n

5.2.1 μ-Ordinary abelian schemes over R_m-algebras

Let $m \geq 1$ and let R be an R_m-algebra. If $\underline{A} \in S_\mu^{(m)}(R) \subset \mathcal{M}(R)$, then A is fiber-by-fiber μ-ordinary, hence $A[p^n]^\mu$, the largest R-subgroup scheme of $A[p^n]$ of multiplicative type (dual to the étale quotient $A[p^n]^{et}$), is a finite flat $\mathcal{O}_\mathcal{K}$-subgroup scheme of rank p^{2n}. Locally in the étale topology it is isomorphic to $\delta_\mathcal{K}^{-1}\mathcal{O}_\mathcal{K} \otimes \mu_{p^n}$.

5.2.2 Igusa level structure of level p^n

Fix $m \geq 1$ and $n \geq 1$ and consider the moduli problem associating with an R_m-algebra R μ-ordinary tuples $\underline{A} \in S_\mu^{(m)}(R)$ together with an isomorphism of finite flat group schemes over R

$$\varepsilon = \varepsilon_n^{(m)} : \delta_\mathcal{K}^{-1}\mathcal{O}_\mathcal{K} \otimes \mu_{p^n} \simeq A[p^n]^\mu. \tag{5.5}$$

This moduli problem is representable by a scheme $Ig(p^n)_\mu^{(m)}$, and the map "forget ε" is a finite étale cover

$$\tau = \tau_n^{(m)} : Ig(p^n)_\mu^{(m)} \to S_\mu^{(m)} \tag{5.6}$$

of degree $(p^2-1)p^{2(n-1)}$. It extends to a finite étale cover $\overline{Ig}(p^n)_\mu^{(m)}$ of $\bar{S}_\mu^{(m)}$. The group

$$\Delta(p^n) = (\mathcal{O}_\mathcal{K}/p^n\mathcal{O}_\mathcal{K})^\times = Aut_{\mathcal{O}_\mathcal{K}}\left(\delta_\mathcal{K}^{-1}\mathcal{O}_\mathcal{K} \otimes \mu_{p^n}\right) \tag{5.7}$$

acts on the covering τ as a group of deck transformations via

$$\gamma(\underline{A}, \varepsilon) = (\underline{A}, \varepsilon \circ \gamma^{-1}), \tag{5.8}$$

and the pre-image of the cuspidal divisor C is non-canonically isomorphic to $\Delta(p^n) \times C$. These constructions satisfy the usual compatibilities in m and n.

5.2.3 The trivialization of $\mathcal{L}$ when $m \leq n$

Assume now that $m \leq n$. In this case, multiplication by p^n is 0 on R, so the inclusion of $A[p^n]$ in A induces an isomorphism between the cotangent spaces at the origin $\omega_{A[p^n]/R}$ and $\omega_{A/R}$. To see it note that if $\mathcal{G}$ is either $A[p^n]$ or A, its Lie algebra, by definition, is the finite flat R-module

$$Lie(\mathcal{G}) = \ker\left(\mathcal{G}(R[\epsilon]) \to \mathcal{G}(R)\right). \tag{5.9}$$

Here $R[\epsilon]$ is the ring of dual numbers over R. It follows that

$$Lie(A[p^n]) = Lie(A)[p^n] = Lie(A), \tag{5.10}$$

and dualizing we get $\omega_{A/R} = \omega_{A[p^n]/R}$.

The same holds of course for μ_{p^n} and $\mathbb{G}_m$. The reasoning used for $m = n = 1$ applies and shows that ε induces a canonical isomorphism between $\mathcal{L}|_{\overline{Ig}(p^n)_\mu^{(m)}}$ and $\mathcal{O}_{\overline{Ig}(p^n)_\mu^{(m)}}$. We denote by $a = a_n^{(m)}$ the section which corresponds to $1 \in \mathcal{O}_{\overline{Ig}(p^n)_\mu^{(m)}}$, i.e. the trivializing section.

The group $\Delta(p^n)$ acts on a via the character

$$\bar{\Sigma}^{-1} : \Delta(p^n) = (\mathcal{O}_\mathcal{K}/p^n\mathcal{O}_\mathcal{K})^\times \to (\mathcal{O}_\mathcal{K}/p^m\mathcal{O}_\mathcal{K})^\times = R_m^\times. \tag{5.11}$$

From now on we take $n = m$ and use a to trivialize $\mathcal{L}$ along $\tilde{C} = \tau^{-1}(C)$, the cuspidal divisor in $\overline{Ig}(p^m)^{(m)}_\mu$. If $f \in H^0(\bar{S}^{(m)}_\mu, \mathcal{L}^k)$, then $\tau^* f/a^k$ is a function on $\overline{Ig}(p^m)^{(m)}_\mu$ and we may attach to it a *canonical* FJ expansion

$$\widetilde{FJ}(f) = \sum_{j=0}^{\infty} c_j(f) \tag{5.12}$$

where $c_j(f) \in H^0(\tilde{C}, \mathcal{N}^j)$ as before. This FJ expansion does not depend on any choice (but is defined along $\tilde{C}$ and not along C).

5.2.4 Congruences between FJ expansions force congruences between the weights

Let $k_1 \leq k_2$ be two integers. The following lemma follows formally from the definitions.

Lemma 5.3 *Let $f_i \in H^0(\bar{S}^{(m)}_\mu, \mathcal{L}^{k_i})$ and assume that f_1 is not divisible by p. Suppose $\widetilde{FJ}(f_1) = \widetilde{FJ}(f_2)$. Then $k_1 \equiv k_2 \mod (p^2-1)p^{m-1}$.*

Proof Let $\tilde{T}$ be an irreducible component of $\overline{Ig}(p^m)^{(m)}_\mu$. Then, τ being finite étale, $\tau(\tilde{T})$ is both open and closed in $\bar{S}^{(m)}_\mu$, so must be an irreducible component T of $\bar{S}^{(m)}_\mu$. It follows that $\tau(\tilde{T})$ meets C, hence $\tilde{T}$ meets $\tilde{C}$, and the q-expansion principle holds in $\overline{Ig}(p^m)^{(m)}_\mu$. We therefore have an equality

$$\tau^* f_1/a^{k_1} = \tau^* f_2/a^{k_2} \tag{5.13}$$

between functions on $\overline{Ig}(p^m)^{(m)}_\mu$. Since the left-hand side is not divisible by p by assumption, so is the right-hand side. The group $\Delta(p^m)$ acts on the left-hand side via $\bar{\Sigma}^{k_1}$ and on the right-hand side via $\bar{\Sigma}^{k_2}$. But these two characters are equal if and only if $k_1 \equiv k_2 \mod (p^2-1)p^{m-1}$, because the exponent of the group $\Delta(p^m)$ is $(p^2-1)p^{m-1}$.

In practice, one would like to deduce the same result from congruences between FJ expansions along C, not along $\tilde{C}$. This is deeper and depends on Igusa's irreducibility theorem.

Theorem 5.4 *Consider $\tau = \tau^{(1)}_n : \overline{Ig}(p^n)^{(1)}_\mu \to \bar{S}^{(1)}_\mu = \bar{S}_{\mu,\kappa_0}$ and extend scalars from κ_0 to κ. Let T be an irreducible component of $\bar{S}_{\mu,\kappa}$. Then $\tau^{-1}(T)$ is irreducible in $\overline{Ig}(p^n)_{\mu,\kappa}$.*

Proof The theorem can be proved by the same method used by Hida [16, 8.4], [17], or by the method of Ribet to which we alluded in 2.4.3. In that section, we proved the theorem for $n = 1$ by a third method, due to Igusa, studying the image of inertia around S_{ss}. See also the discussion of the big Igusa tower *BigIg* below, which turns out to be reducible.

Theorem 5.5 *Let f_1 and f_2 be $\mod p^m$ modular forms as above and assume that f_1 is not divisible by p. Trivialize $\mathcal{L}|_C$ by choosing a lift of C to $\tilde{C}$ (i.e. a section of the map $\tau|_{\tilde{C}} : \tilde{C} \to C$) and using the trivialization of $\mathcal{L}$ along this lift which is supplied by the section a. Then if $FJ(f_1) = FJ(f_2)$, $k_1 \equiv k_2 \mod (p^2-1)p^{m-1}$.*

Here $FJ(f) = \sum_{j=0}^{\infty} c_j(f)$ and $c_j(f) \in H^0(C, \mathcal{N}^j)$. The lift of C to $\tilde{C}$ exists since $\tilde{C} \simeq \Delta(p^m) \times C$ (non-canonically). If we change the lift (locally on the base) by $\gamma \in \Delta(p^m)$, then $FJ(f_i)$ changes by the factor $\bar{\Sigma}(\gamma)^{k_i}$.

Proof By Igusa's irreducibility theorem, it is enough to know that $FJ(f_i)$ $(i = 1, 2)$ agree on the given lift of C, to conclude that $\tau^* f_1/a^{k_1} = \tau^* f_2/a^{k_2}$ on the whole of $\overline{Ig}(p^m)^{(m)}_\mu$, hence the result follows by the Lemma. Note that the underlying topological spaces of $\overline{Ig}(p^m)_\mu^{(m)}$ and $\overline{Ig}(p^m)_\mu^{(1)}$ are the same, and hence for the irreducibility theorem, it is enough to deal with the special fiber.

Corollary 5.6 *Let $f_i \in M^p_{k_i}(N; \mathcal{O}_p)$ $(i = 1, 2)$ and assume that f_1 is not divisible by p. Trivialize $\mathcal{L}|_C$ by fixing an $\mathcal{O}_\mathcal{K}$-isomorphism of the p-divisible group of the toric part of the universal semi-abelian variety $\mathcal{A}|_C$ with $\delta_\mathcal{K}^{-1}\mathcal{O}_\mathcal{K} \otimes \mu_{p^\infty}$, and using this isomorphism to identify $\mathcal{L}|_C = \omega_{\mathcal{A}/C}(\bar{\Sigma})$ with $\mathcal{O}_C$. Suppose that with this trivialization*

$$FJ(f_1) \equiv FJ(f_2) \mod p^m. \tag{5.14}$$

Then $k_1 \equiv k_2 \mod \left(p^2 - 1\right) p^{m-1}$.

5.2.5 Irreducibility of the Igusa tower and the big Igusa tower

It is possible to define an even larger Igusa tower $(BigIg(p^n))_{n\geq 1}$ over $\kappa = \bar{\mathbb{F}}_p$, of which $(Ig(p^n))_{n\geq 1}$ is a quotient. If R is a κ-algebra and $\underline{A} \in S_\mu(R)$, then $A[p^n]$ admits a filtration as in 2.1.2. One can define $BigIg(p^n)$ as the moduli space of μ-ordinary tuples $\underline{A}$, equipped with $\mathcal{O}_\mathcal{K}$-isomorphisms

$$\begin{aligned} &\varepsilon^2 : \delta_\mathcal{K}^{-1}\mathcal{O}_K \otimes \mu_{p^n} \simeq gr^2 A[p^n] \\ &\varepsilon^1 : \mathfrak{G}[p^n] \simeq gr^1 A[p^n] \\ &\varepsilon^0 : \mathcal{O}_\mathcal{K} \otimes \mathbb{Z}/p^n\mathbb{Z} \simeq gr^0 A[p^n]. \end{aligned} \tag{5.15}$$

This would be, in the language of [17], the GU-Igusa tower. If we insist that the isomorphisms respect the pairings induced on these group schemes by the polarization and Cartier duality (gr^0 and gr^2 are dual to each other, gr^1 is self-dual), we would get the U-Igusa tower. Both these towers are reducible, by the reasoning of [16, 8.4.1] or [17], and by the description of the connected components of the characteristic 0 fiber of the Shimura variety given in 1.3.3. The SU-Igusa tower, which is *irreducible*, turns out to be our tower $(Ig(p^n))$. It is also the quotient of $(BigIg(p^n))$ under the map "forget ε^0 and ε^1". Thus there is no real advantage in studying the tower *BigIg*.

5.3 *p*-adic modular forms of *p*-adic weights

5.3.1 The space of p-adic weights

Let

$$\mathfrak{X}_p = \varprojlim \mathbb{Z}/(p^2-1)p^{m-1}\mathbb{Z}, \tag{5.16}$$

This is the space of p-adic weights. If $k \in \mathfrak{X}_p$, then $\bar{\Sigma}^k$ is a well-defined locally $\mathbb{Q}_p$-analytic homomorphism of $\mathcal{O}_p^\times$ to itself, but note that not every such homomorphism is a $\bar{\Sigma}^k$ for some k from $\mathfrak{X}_p$.

5.3.2 p-adic modular forms à la Serre

We work with $\bar{S}$ (hence also the cuspidal divisor C) over the base $\mathcal{O}_p$, the p-adic completion of R_0. Little is lost by extending the base further to $\mathcal{O}_{N,\mathfrak{P}}$, the completion of the ring of integers of the ray class field $\mathcal{K}_N$ at a prime $\mathfrak{P}$ above p. After such a base extension the irreducible components of C become absolutely irreducible. The reader may assume that this is the case.

Consider the p-divisible group of the toric part of the universal semi-abelian variety $\mathcal{A}|_C$. Once and for all fix an $\mathcal{O}_\mathcal{K}$-isomorphism of it with $\delta_\mathcal{K}^{-1}\mathcal{O}_\mathcal{K} \otimes \mu_{p^\infty}$, and use this isomorphism to identify $\mathcal{L}|_C = \omega_{\mathcal{A}/C}(\bar{\Sigma})$ with $\mathcal{O}_C$. This choice is unique up to multiplication by $\mathcal{O}_p^\times$ on each irreducible component of C. It determines a FJ expansion $FJ(f)$ for every $f \in M_k^p(N;\mathcal{K}_p)$ as in (1.125), and is equivalent to splitting the projection $\tau|_{\tilde{C}} : \tilde{C} \to C$ from the boundary of the Igusa tower $\left(\overline{Ig}_\mu(p^n)\right)_{n=1}^\infty$ to the boundary of the Picard modular surface.

Let $k \in \mathfrak{X}_p$. The space $M_k^{\mathrm{Serre}}(N;\mathcal{K}_p)$ will be a subspace of the Banach algebra

$$\mathcal{FJ}_p = \mathcal{K}_p \otimes_{\mathcal{O}_p} \prod_{j=0}^\infty H^0(C, \mathcal{N}^j). \tag{5.17}$$

It will consist of all the $f \in \mathcal{FJ}_p$ for which there exists a sequence (f_ν), $f_\nu \in M_{k_\nu}^p(N;\mathcal{K}_p)$, $(k_\nu \in \mathbb{Z})$, with $FJ(f_\nu)$ converging to f, and k_ν converging in $\mathfrak{X}_p$ to k. As we have seen, if the sequence $(FJ(f_\nu))$ converges, the k_ν have to converge in $\mathfrak{X}_p$. We shall denote by $M_k^{\mathrm{Serre}}(N;\mathcal{O}_p)$ the intersection of $M_k^{\mathrm{Serre}}(N;\mathcal{K}_p)$ with $\prod_{j=0}^\infty H^0(C, \mathcal{N}^j)$.

Proposition 5.7 (i) *If $k \in \mathbb{Z}$, then $M_k^{\mathrm{Serre}}(N;\mathcal{K}_p) = M_k^p(N;\mathcal{K}_p)$. In other words, we do not get any new p-adic modular forms by allowing limits of p-adic modular forms of varying weights, if the weights converge to an integral k.*

(ii) *In the definition of $M_k^{\mathrm{Serre}}(N;\mathcal{K}_p)$ we can require $f_\nu \in M_{k_\nu}(N;\mathcal{K}_p)$ (classical modular forms of integral weight k_ν) and still get the same space.*

(iii) *$M_k^{\mathrm{Serre}}(N;\mathcal{K}_p)$ is a closed subspace of $\mathcal{FJ}_p$. The product of two $f_i \in M_{k_i}^{\mathrm{Serre}}$ is in $M_{k_1+k_2}^{\mathrm{Serre}}$.*

(iv) *If $f \in M_k^{\mathrm{Serre}}(N;\mathcal{O}_p)$, then its reduction modulo p appears in $M_{k'}(N;\kappa_0)$ for some positive integer k' sufficiently close to k in $\mathfrak{X}_p$.*

Proof Let $H_{\bar{\Sigma}} \in M_{p^2-1}(N;\mathcal{O}_p)$ be a lift of the Hasse invariant $h_{\bar{\Sigma}}$ to characteristic 0. Such a lift exists by general principles, whenever p is large enough. For the few exceptional primes p we may replace $h_{\mathbf{\Sigma}}$ by a high enough power of it, which is liftable, and use the same argument. This lift satisfies $FJ(H_{\bar{\Sigma}}) \equiv 1 \mod p$, so $H_{\bar{\Sigma}}^{-1} \in M_{1-p^2}^p(N;\mathcal{O}_p)$ is a p-adic modular form defined over $\mathcal{O}_p$. Indeed, $H_{\bar{\Sigma}} \mod p^m \in H^0(\bar{S}_\mu^{(m)}, \mathcal{L}^{p^2-1})$ is nowhere vanishing over $\bar{S}_\mu^{(m)}$, and taking the limit of its inverse over m we get $H_{\bar{\Sigma}}^{-1}$. Suppose, as in (i), that $k, k_\nu \in \mathbb{Z}$, $k_\nu \to k$ in $\mathfrak{X}_p$, and $f_\nu \in M_{k_\nu}^p(N;\mathcal{K}_p)$ are such that $FJ(f_\nu)$ converge in $\mathcal{FJ}_p$ to f. Replacing f_ν by $f_\nu H_{\bar{\Sigma}}^{p^{e_\nu}}$ for suitable e_ν, we may assume that the k_ν are increasing and are all in the same congruence class modulo $p^2 - 1$. But then $f_\nu H_{\bar{\Sigma}}^{(k-k_\nu)/(p^2-1)}$ are in $M_k^p(N;\mathcal{K}_p)$ and their FJ expansions converge to f in $\mathcal{FJ}_p$. This proves (i). For (ii) note that if $f \in H^0(\bar{S}_\mu^{(m)}, \mathcal{L}^k)$, then for all sufficiently large e, $fH_{\bar{\Sigma}}^{p^e}$ extends to an element of $M_{k+(p^2-1)p^e}(N;R_m)$ and has the same FJ expansion as f. Thus every p-adic modular form of integral weight is the p-adic limit of classical forms of varying weights, and the same is therefore true for Serre modular forms of p-adic weight. Points (iii) and (iv) are obvious.

5.3.3 p-Adic modular forms à la Katz

We now explain Katz' point of view of the same objects. Let

$$V_n^{(m)} = H^0\left(\overline{Ig}(p^n)_\mu^{(m)}, \mathcal{O}\right) \tag{5.18}$$

be the ring of regular functions on $\overline{Ig}(p^n)_\mu^{(m)}$. Let

$$V^{(m)} = \varinjlim V_n^{(m)}, \qquad V = \varprojlim V^{(m)}. \tag{5.19}$$

We call V the space of *Katz p-adic modular forms* (of all weights). Let

$$\gamma \in \Delta = \mathcal{O}_p^\times = \varprojlim \Delta(p^n) \tag{5.20}$$

act on $V^{(m)}$ and on V as usual, $\gamma(f) = f \circ \gamma^{-1}$, and recall that $\gamma^{-1}(\underline{A}, \varepsilon_n^{(m)}) = (\underline{A}, \varepsilon_n^{(m)} \circ \gamma)$. Thus

$$\gamma(f)(\underline{A}, \varepsilon) = f(\underline{A}, \varepsilon \circ \gamma) \tag{5.21}$$

(i.e. γ acts by "right translation"). Let $k \in \mathfrak{X}_p$ and define

$$M_k^{\mathrm{Katz}}(N; \mathcal{O}_p) = V(\bar{\Sigma}^k) = \left\{ f \in V \,|\, \gamma(f) = \bar{\Sigma}^k(\gamma) \cdot f \quad \forall \gamma \in \Delta \right\}. \tag{5.22}$$

We similarly define $M_k^{\mathrm{Katz}}(N; R_m) = V^{(m)}(\bar{\Sigma}^k)$.

By the irreducibility of the Igusa tower and the q-expansion principle the FJ expansion map

$$V \to \mathcal{FJ}_p(\mathcal{O}_p) \tag{5.23}$$

is injective. It depends on our choice of the splitting of $\tilde{C} \to C$.

Proposition 5.8 *For $k \in \mathfrak{X}_p$, there is a natural isomorphism*

$$M_k^{\mathrm{Serre}}(N; \mathcal{O}_p) \simeq M_k^{\mathrm{Katz}}(N; \mathcal{O}_p). \tag{5.24}$$

Proof Given $k \in \mathbb{Z}$ and $f \in H^0(\bar{S}_\mu^{(m)}, \mathcal{L}^k)$, the functions $(\tau_n^{(m)})^* f / (a_n^{(m)})^k \in V_n^{(m)}$ for all $n \geq m$, and these functions satisfy the obvious compatibility in n, so they define

$$f^{\mathrm{Katz}} \in V^{(m)}(\bar{\Sigma}^k). \tag{5.25}$$

If $k \in \mathbb{Z}$, this gives, by going to the inverse limit over m, a map

$$f \mapsto f^{\mathrm{Katz}}, \quad M_k^p(N; \mathcal{O}_p) \to M_k^{\mathrm{Katz}}(N; \mathcal{O}_p). \tag{5.26}$$

This map is an isomorphism, which can be enhanced to include p-adic weights $k \in \mathfrak{X}_p$ as follows. If $k_\nu \in \mathbb{Z}$, $k_\nu \to k \in \mathfrak{X}_p$ and if $f_\nu \in M_{k_\nu}^p(N; \mathcal{O}_p)$ are such that $FJ(f_\nu)$ converge to $f \in M_k^{\mathrm{Serre}}(N; \mathcal{O}_p)$, then reducing modulo p^m for a fixed m, $(f_\nu^{(m)})^{\mathrm{Katz}} \in V^{(m)}(\bar{\Sigma}^{k_\nu})$. But for a fixed m, for all large enough ν,

$$V^{(m)}(\bar{\Sigma}^{k_\nu}) = V^{(m)}(\bar{\Sigma}^k), \tag{5.27}$$

and the sequence $FJ(f_\nu^{(m)})$ stabilizes, so taking the limit over ν we get a well defined $(f^{(m)})^{\mathrm{Katz}} \in V^{(m)}(\bar{\Sigma}^k)$. Finally, an inverse limit over m gives $f^{\mathrm{Katz}} \in M_k^{\mathrm{Katz}}(N; \mathcal{O}_p)$. It is by now standard that this gives an isomorphism between $M_k^{\mathrm{Serre}}(N; \mathcal{O}_p)$ and $M_k^{\mathrm{Katz}}(N; \mathcal{O}_p)$. As we have seen earlier, when $k \in \mathbb{Z}$, this is also the same as $M_k^p(N; \mathcal{O}_p)$.

From now on, it is therefore legitimate to denote these spaces by the common notation $M_k^p(N; \mathcal{O}_p)$ and refer to them simply as *p-adic modular forms* of *p-adic weight k.*

5.4 *p*-Adic modular forms of *p*-adic bi-weights

5.4.1 The space of bi-weights

A new feature of p-adic modular forms on Picard modular surfaces, that does not show up in the classical theory of $GL_2(\mathbb{Q})$, is that even if we restrict attention to scalar-valued p-adic modular forms, we sometimes need to consider classical *vector-valued* forms to approach them. This phenomenon, as we shall explain below, does not show up in the mod p theory, but is essential to the p-adic theory.

The space $\mathfrak{X}_p$ of p-adic weights can be written as $\mathbb{Z}/(p^2-1)\mathbb{Z} \times \mathbb{Z}_p$, and when we decompose it in such a way we write

$$k = (w, j) = \big(\omega(k), \langle k \rangle\big) \tag{5.28}$$

for the two components. The space of *bi-weights* $\mathfrak{X}_p^{(2)}$ is, by definition, the quotient of $\mathfrak{X}_p^2$ modulo the relation

$$((w_1, j_1), (w_2, j_2)) \equiv ((0, j_1), (pw_1 + w_2, j_2)) \equiv ((pw_2 + w_1, j_1), (0, j_2)). \tag{5.29}$$

If k_1 and k_2 are in $\mathfrak{X}_p$, then the character $\bar{\Sigma}^{k_1}\Sigma^{k_2} : \Delta \to \mathcal{O}_p^\times$ depends only on the image of (k_1, k_2) in $\mathfrak{X}_p^{(2)}$. Here $\Delta = \varprojlim \Delta(p^n)$ is also $\mathcal{O}_p^\times$, but in the rôle of the Galois group of the Igusa tower. The image of $\mathbb{Z}^2$ is dense in $\mathfrak{X}_p^2$, hence also in $\mathfrak{X}_p^{(2)}$.

5.4.2 The line bundle $\mathcal{L}^{(k_1,k_2)}$ over $\bar{S}_\mu^{rig}$ and p-adic modular forms of integral bi-weights

Let $m \geq 1$. The plane bundle $\mathcal{P}$ admits a canonical filtration

$$0 \to \mathcal{P}_0 \to \mathcal{P} \to \mathcal{P}_\mu \to 0 \tag{5.30}$$

over $\bar{S}_\mu^{(m)}$ defined by choosing any $n \geq m$ and setting

$$\mathcal{P}_0 = \ker(\omega_{\mathcal{A}[p^n]^0} \to \omega_{\mathcal{A}[p^n]^\mu}),\ \ \mathcal{P}_\mu = \omega_{\mathcal{A}[p^n]^\mu}(\Sigma)$$

(recall $\omega_{\mathcal{A}} = \omega_{\mathcal{A}[p^n]^0}$). We also recall that $\mathcal{L} = \omega_{\mathcal{A}[p^n]^\mu}(\bar{\Sigma})$.

If $m = 1$ we showed that over $\bar{S}_\mu^{(1)}$, $\mathcal{L} \simeq \mathcal{P}_\mu^p$ and $\mathcal{P}_\mu \simeq \mathcal{L}^p$. This is no longer true for general m and we let for $(k_1, k_2) \in \mathbb{Z}^2$

$$\mathcal{L}^{(k_1,k_2)} = \mathcal{L}^{k_1} \otimes \mathcal{P}_\mu^{k_2}. \tag{5.31}$$

Going to the limit over m, this defines a rigid analytic line bundle over $\bar{S}_\mu^{rig}$.

We define the space of *p-adic modular forms of bi-weight* (k_1, k_2) *and level* N over $\mathcal{O}_p$ as

$$M_{k_1,k_2}^p(N; \mathcal{O}_p) = \varprojlim H^0\left(\bar{S}_\mu^{(m)},\ \mathcal{L}^{(k_1,k_2)}\right). \tag{5.32}$$

This is the unit ball of the p-adic Banach space

$$M_{k_1,k_2}^p(N; \mathcal{K}_p) = \mathcal{K}_p \otimes_{\mathcal{O}_p} M_{k_1,k_2}^p(N; \mathcal{O}_p) = H^0\left(\bar{S}_\mu^{rig}, \mathcal{L}^{(k_1,k_2)}\right). \tag{5.33}$$

5.4.3 The trivialization of $\mathcal{L}^{(k_1,k_2)}$ over the Igusa tower

As before, fix m, let $m \leq n$ and consider the isomorphism

$$\varepsilon^* : \tau^*\omega_{\mathcal{A}[p^n]^\mu} \simeq \mathcal{O}_{\mathcal{K}} \otimes \mathcal{O}_{\overline{Ig}(p^n)_\mu^{(m)}} \tag{5.34}$$

induced by the Igusa level structure $\varepsilon = \varepsilon_n^{(m)}$. Taking $\bar{\Sigma}$ and Σ-types, it induces trivializations

$$\tau^*\mathcal{L} \simeq \mathcal{O}_{\overline{Ig}(p^n)_\mu^{(m)}}, \quad \tau^*\mathcal{P}_\mu \simeq \mathcal{O}_{\overline{Ig}(p^n)_\mu^{(m)}} \tag{5.35}$$

and we let $a = a_n^{(m)}$ and $\bar{a} = \bar{a}_n^{(m)}$ be the sections corresponding to 1. Of course, the trivialization of $\mathcal{L}$ is the one that we have met before.

Let $a^{k_1,k_2} = a^{k_1}\bar{a}^{k_2}$. Then we may trivialize $\tau^*\mathcal{L}^{(k_1,k_2)}$ by $s \mapsto s/a^{k_1,k_2}$ to get a function on $\overline{Ig}(p^n)^{(m)}_\mu$. This allows us to define, as usual, canonical Fourier–Jacobi expansion $\widetilde{FJ}(f)$ (along $\tilde{C}$), and if we make a choice of a splitting of $\tau : \tilde{C} \to C$, a Fourier–Jacobi expansion $FJ(f)$ (along C) for every $f \in M^p_{k_1,k_2}(N;\mathcal{K}_p)$.

5.4.4 p-Adic modular forms of p-adic bi-weights

The yoga of p-adic weights, either à la Serre or à la Katz, allows us now to define the space

$$M^p_{k_1,k_2}(N;\mathcal{K}_p) \tag{5.36}$$

of p-adic modular forms of any bi-weight $(k_1,k_2) \in \mathfrak{X}^{(2)}_p$. If we follow Serre, we define them as elements of the Banach space $\mathcal{FJ}_p$ via limits of p-adic modular forms of integral bi-weights. If we follow Katz, we have

$$M^p_{k_1,k_2}(N;\mathcal{O}_p) = V\left(\bar{\Sigma}^{k_1}\Sigma^{k_2}\right). \tag{5.37}$$

We let the reader complete the details, which are identical to the case of a single weight treated before.

5.5 The theta operator for *p*-adic modular forms

We are finally able to define the operator Θ on p-adic modular forms. Compare [22, V.5.8]. Let $f \in M^p_{k_1,k_2}(N;\mathcal{O}_p)$. Assume first that k_1 and k_2 are from $\mathbb{Z}$, and reduce modulo p^m, to get $f \in H^0(\bar{S}^{(m)}_\mu, \mathcal{L}^{k_1} \otimes \mathcal{P}^{k_2}_\mu)$. Take any $n \geq m$, pull back to $\overline{Ig}(p^n)^{(m)}_\mu$, divide by a^{k_1,k_2} and consider

$$\eta_f = \mathrm{d}\left(\tau^* f/a^{k_1,k_2}\right) \in H^0\left(\overline{Ig}(p^n)^{(m)}_\mu, \Omega^1_{Ig}\right). \tag{5.38}$$

Apply KS^{-1} to η_f. This results in a section of $\mathcal{L}\otimes\mathcal{P}$. As explained before, when we project this section to $\mathcal{L}\otimes\mathcal{P}_\mu$, we get a section that is holomorphic along $\tilde{C}$ and even vanishes there (recall KS had a pole along the cuspidal divisor). Multiply back by a_{k_1,k_2} and use Galois descent to descend the resulting section to $S^{(m)}_\mu$.

We may now take the limit over m to get our Θ, if $(k_1,k_2) \in \mathbb{Z}^2$. A further limit over weights, as in the proof of Proposition 5.8, allows us to extend the definition to $(k_1,k_2) \in \mathfrak{X}^{(2)}_p$. Using Katz' approach, where the process of dividing and multiplying back by a^{k_1,k_2} is already built into the isomorphism with $V(\bar{\Sigma}^{k_1}\Sigma^{k_2})$, Θ is nothing but the map

$$\Theta : f \mapsto (1\otimes pr_\mu)\circ \mathrm{KS}^{-1}\circ d(f) \tag{5.39}$$

sending $V(\bar{\Sigma}^{k_1}\Sigma^{k_2})$ to $V(\bar{\Sigma}^{k_1+1}\Sigma^{k_2+1})$. Here $pr_\mu : \mathcal{P} \to \mathcal{P}_\mu$ is the projection defined over $\bar{S}^{rig}_\mu$.

Theorem 5.9 *Let* $(k_1,k_2) \in \mathfrak{X}^{(2)}_p$. *The operator,*

$$\Theta : M^p_{k_1,k_2}(N;\mathcal{O}_p) \to M^p_{k_1+1,k_2+1}(N;\mathcal{O}_p) \tag{5.40}$$

defined by the above formula, satisfies the following properties (and is uniquely determined by its effect on q-expansions).

(i) *When one reduces* $M^p_{k_1,k_2}(N;\mathcal{O}_p)$ *modulo p, and uses the isomorphism* $\mathcal{P}_\mu \simeq \mathcal{L}^p$, Θ *reduces to the operator*

$$\Theta : M_k(N;\kappa) \to M_{k+p+1}(N;\kappa) \tag{5.41}$$

on mod p *modular forms.*

(ii) *The effect of* Θ *on the canonical FJ expansion* $\widetilde{FJ}(f)$ *is given by "*$q\frac{d}{dq}$*", i.e. by the formula* (3.21).

We omit the proof of (ii), which goes along the same lines as in the mod p theory.

Author details

[1]Hebrew University, Jerusalem, Israel , [2]McGill University, Montreal, Canada.

Received: 5 July 2015 Accepted: 27 May 2016
Published online: 01 September 2016

References

1. Andreatta, F., Goren, E.Z.: Hilbert modular forms: mod *p* and *p*-adic aspects, Memoirs A.M.S. **819** (2005)
2. Bellaïche, J.: Congruences endoscopiques et représentations Galoisiennes. Thèse, Paris XI (Orsay) (2002)
3. Böcherer, S., Nagaoka, S.: On mod p properties of Siegel modular forms. Math. Ann. **338**, 421–433 (2007)
4. Bültel, O., Wedhorn, T.: Congruence relations for Shimura varieties associated to some unitary groups. J. Inst. Math. Jussieu **5**, 229–261 (2006)
5. Cogdell, J.: Arithmetic cycles on Picard modular surfaces and modular forms of Nebentypus. J. Reine Angew. Math. **357**, 115–137 (1985)
6. Coleman, R.: Classical and overconvergent modular forms. Invent. Math. **124**, 215–241 (1996)
7. Deligne, P.: Travaux de Shimura. Sém. Bourbaki **389**, 123–165 (1971)
8. de Shalit, E., Goren, E.Z.: Supersingular curves on Picard modular surfaces modulo an inert prime. J. Number Theor. (2015) (to appear)
9. Eischen, E.: *p*-Adic differential operators on automorphic forms on unitary groups. Annales de l'Institut Fourier **62**, 177–243 (2012)
10. Eischen, E.: A *p*-adic Eisenstein measure for unitary groups. J. Reine Angew. Math. **699**, 111–142 (2015)
11. Faltings, G., Chai, C.-L.: Degeneration of Abelian Varieties. Springer, Berlin (1990)
12. Goren, E.Z.: Hilbert modular forms modulo p^m—the unramified case. J. Number Theory **90**, 341–375 (2001)
13. Gross, B.: A tameness criterion for Galois representations associated to modular forms (mod *p*). Duke Math. J. **61**, 445–517 (1990)
14. Harris, M.: Arithmetic vector bundles and automorphic forms on Shimura varieties. I. Invent. Math. **82**, 151–189 (1985)
15. Hartshorne, R.: Algebraic Geometry, Graduate Texts in Mathematics, No. 52. Springer, Berlin (1977)
16. Hida, H.: *p*-Adic Automorphic Forms on Shimura Varieties. Springer, Berlin (2004)
17. Hida, H.: Irreducibility of the Igusa tower. Acta Math. Sin. Engl. Ser. **25**, 1–20 (2009)
18. Hsieh, M.-L.: Eisenstein congruence on unitary groups and Iwasawa main conjectures for CM fields. J. AMS **27**, 753–862 (2014)
19. Jochnowitz, N.: A study of the local components of the Hecke algebra mod *l*. Trans. Am. Math. Soc. **270**, 253–267 (1982)
20. Katz, N.: *p*-Adic properties of modular schemes and modular forms. In: Serre, K. (ed.) Modular Functions of One Variable III. LNM vol. 350, pp. 69–190. Springer, Berlin (1973)
21. Katz, N.: A result on modular forms in characteristic *p*. In: Zagier, S. (ed.) Modular Functions of One Variable V. LNM vol. 601, pp. 53–61. Springer, Berlin (1977)
22. Katz, N.: *p*-Adic interpolation of real analytic Eisenstein series. Ann. Math. **104**, 459–571 (1976)
23. Katz, N., Mazur, B.: Arithmetic moduli of elliptic curves. Ann. Math. Stud. **108** (1985)
24. Katz, N., Oda, T.: On the differentiation of de Rham cohomology classes with respect to parameters. J. Math. Kyoto Univ. **8**, 199–213 (1968)
25. Koblitz, N.: *p*-Adic variation of the zeta-function over families of varieties defined over finite fields. Compos. Math. **31**, 119–218 (1975)
26. Lan, K.-W.: Arithmetic Compactifications of PEL-Type Shimura Varieties, London Mathematical Society Monographs vol. 36. Princeton University Press, Princeton (2013)
27. Langlands, R., Ramakrishnan, D.: The Zeta Functions of Picard Modular Surfaces. Univ. Montréal, Montréal (1992)
28. Larsen, M.: Unitary groups and *l*-adic representations, Ph.D. Thesis, Princeton University (1988)
29. Larsen, M.: Arithmetic compactification of some Shimura surfaces. In: [28], pp. 31–45
30. Moonen, B.: Group schemes with additional structures and Weyl group cosets. In: van der Geer Oort, F. (ed). Moduli of Abelian Varieties. Progress in Mathematics, vol. 195, pp. 255–298. Birkhäuser, Basel (2001)
31. Mumford, D.: Abelian Varieties. Oxford University Press, London (1970)
32. Mumford, D.: The Red Book of Varieties and Schemes, Second Expanded Edition, LNM, vol. 1358. Springer, Berlin (1999)
33. Ribet, K.: A *p*-adic interpolation via Hilbert modular forms. In: Algebraic Geometry. Proceedings of Symposia in Pure Mathematics, vol. 29, pp. 581–592. AMS (1975)
34. Serre, J.-P.: Formes modulaires et fonctions zêta *p*-adiques. In: Serre, K. (ed.) Modular Functions of One Variable III, LNM, vol. 350, pp. 191–268. Springer, Berlin (1973)
35. Shimura, G.: Arithmetic of unitary groups. Ann. Math. **79**, 369–409 (1964)
36. Shimura, G.: The arithmetic of automorphic forms with respect to a unitary group. Ann. Math. **107**, 569–605 (1978)
37. Shimura, G.: Arithmeticity in the Theory of Automorphic Forms. Mathematical Surveys and Monographs, vol. 82. AMS, Providence (2000)

38. Swinnerton-Dyer, H.P.F.: On l-adic representations and congruences for coefficients of modular forms. In: Serre, K. (ed.) Modular Functions of One Variable III, LNM, vol. 350, pp. 1–55. Springer, Berlin (1973)
39. Vollaard, I.: The supersingular locus of the Shimura variety for $GU(1, s)$. Can. J. Math. **62**, 668–720 (2010)
40. Wedhorn, T.: The dimension of Oort strata of Shimura varieties of PEL-type. In: van der Geer Oort, F. (ed). Moduli of Abelian Varieties. Progress in Mathematics, vol. 195, pp. 441–471. Birkhäuser, Basel (2001)