Contents

CHAPTER 1

Tori and Abelian Varieties

In the first section of this chapter we recall the definition of algebraic groups and some basic facts about them, attempting to put the classes of groups that will interest us – abelian varieties and algebraic tori – in a somewhat larger perspective. We quickly specialize to diagonalizable groups while the next sections treat abelian varieties.

There are several good reasons to discuss tori in detail. They appear as degenerations of abelian varieties, their characters are weights of Hilbert modular forms, and their finite subgroups are involved in defining level structures on abelian varieties.

In general, we assume that this is not the first time the reader is exposed to these topics. In particular, we assume familiarity with basic algebraic geometry, elliptic curves and elliptic modular forms.

1. Algebraic Groups

In this section we follow very closely Borel [**5**], where the reader may find also a moderate introduction to algebraic geometry.

Let k be a field with an algebraic closure K and let k^s be the separable closure of k in K. We denote $Gal(k^s/k)$ by Γ throughout this section. We shall denote by $k[G]$ the ring of regular functions on a variety G over the k.

DEFINITION 1.1. An *algebraic group* over k is a quasi-projective variety G over k together with morphisms

$$m : G \times G \longrightarrow G, \tag{1.1}$$

$$inv : G \longrightarrow G, \tag{1.2}$$

and identity element e, such that the following diagrams commute:

$$\begin{array}{ccc} G\times G\times G & \xrightarrow{m\times 1} & G\times G \\ {\scriptstyle 1\times m}\downarrow & & \downarrow{\scriptstyle m,} \\ G\times G & \xrightarrow[m]{} & G \end{array} \qquad \begin{array}{ccc} G & \xrightarrow{(e,1)} & G\times G \\ {\scriptstyle (1,e)}\downarrow & \searrow^{1} & \downarrow{\scriptstyle m,} \\ G\times G & \xrightarrow[m]{} & G \end{array} \qquad \begin{array}{ccc} G & \xrightarrow{(inv,1)} & G\times G \\ {\scriptstyle (1,inv)}\downarrow & \searrow^{e} & \downarrow{\scriptstyle m.} \\ G\times G & \xrightarrow[m]{} & G \end{array} \tag{1.3}$$

Where we use e to denote also the constant map $G \longrightarrow e$.

Thus, from a "classical" point of view, an algebraic group G is nothing else than a group $G(K)$, with identity element $e \in G(k)$, such that the group operations are algebraic. From a "modern" point of view, i.e., scheme-theoretic, an algebraic group is a functor associating to each k-algebra R the group of R-rational points

of G, $G(R)$, such that the natural maps $G(R) \longrightarrow G(S)$ induced from a morphism of k-algebras $R \longrightarrow S$ are all group homomorphisms. Group schemes are discussed in Appendix A.

An algebraic group G is called an *affine* or *linear* algebraic group if it is an affine variety or, equivalently, if it can be embedded as group variety in $\mathrm{GL}_n(K)$ for some n; see [**5**, Prop. 1.10, p.54]. Thus, an affine algebraic group is a subgroup of GL_n defined by a collection of polynomial equations. Examples are provided by $\mathrm{SL}_n, \mathrm{O_n}, \mathrm{SO_n}, \mathrm{Sp_{2n}}$, the upper triangular matrices, the unipotent matrices. Here the reader will notice that he is already familiar with, e.g., the $\iota\delta\epsilon\alpha$ GL_n. That is, with a functor associating to any ring R a group $GL_n(R)$. Thus, GL_n becomes more a "group machine" than a particular group. That is precisely the scheme theoretic point of view.

If a group G is affine, the group structure may be equivalently defined by giving the maps, dual to (1.3), on the coordinate ring:

$$\widetilde{m} : k[G] \longrightarrow k[G \times G] \cong k[G] \otimes_k k[G], \tag{1.4}$$

$$\widetilde{inv} : k[G] \longrightarrow k[G], \tag{1.5}$$

$$\widetilde{e} : k[G] \longrightarrow k. \tag{1.6}$$

The maps $\widetilde{m}, \widetilde{inv}$, and $\widetilde{e}$ are called co-multiplication, co-inverse and co-unit respectively. They have the very special property of being *ring homomorphisms*. The reader is invited to reflect on that. Such a structure is called a Hopf algebra.

EXAMPLE 1.2. Let $\mathbb{G}_m = \mathrm{GL}_1$ be the multiplicative group. We have

$$k[\mathbb{G}_m] = k[x, x^{-1}] \cong k[x, y]/(xy - 1); \tag{1.7}$$

(the second presentation shows it is affine). The co-multiplication, co-inverse and co-unit maps are:

$$\widetilde{m}(x) = x \otimes x, \quad \widetilde{inv}(x) = x^{-1}, \quad \tilde{e}(x) = 1. \tag{1.8}$$

EXAMPLE 1.3. Let $\mathbb{G}_a = \mathbb{A}^1_k$ be the additive group. We have

$$k[\mathbb{G}_a] = k[x]; \tag{1.9}$$

$$\widetilde{m}(x) = x \otimes 1 + 1 \otimes x, \quad \widetilde{inv}(x) = -x, \quad \tilde{e}(x) = 0. \tag{1.10}$$

EXERCISE 1.4. Write GL_2 as an affine variety. Write its co-multiplication, co-inverse and co-unit morphisms. Do the same for the upper-triangular matrices in GL_2.

An algebraic group G is called *abelian* if it is connected and projective (we will show later that every such abelian group is necessarily commutative). The reader is already familiar with elliptic curves. Except for the zero abelian variety, these are the simplest abelian varieties. Indeed, every abelian variety of dimension one is an elliptic curve. This follows, at least in the case of characteristic zero, from Theorem 6.3 below. A product of elliptic curves is an abelian variety. The Jacobian variety of a smooth curve of genus g is an abelian variety of dimension g, and in fact every abelian variety is a homomorphic image of a Jacobian. See [**73**, Section 10]. See Section 6 for more examples.

If G is an algebraic group, then its identity component G_0 is a normal subgroup of finite index.

EXERCISE 1.5. Prove this! What is the identity component of $O_2(\mathbb{R})$? $O_2(\mathbb{C})$?

THEOREM 1.6. (Chevalley's theorem) *Let G be a connected algebraic group. Then G has a* unique *maximal connected affine subgroup L and G/L is an abelian variety:*

$$0 \longrightarrow L \longrightarrow G \longrightarrow G/L \longrightarrow 0. \tag{1.11}$$

PROOF. See [**12**] and [**15**]. □

REMARK 1.7. In Theorem 1.6 abelian varieties and affine groups appear as opposite sides in the "spectrum" of algebraic groups, yet it may happen that the abelian variety G/L degenerates into the simplest kind of an affine group: a torus (see below). Such degenerations arise for example in the context of compactifications of moduli spaces of abelian varieties.

A typical example of degeneration is the family of elliptic curves over the open unit disk in $\mathbb{C}$:

$$y^2 = x(x-\epsilon)(x-5), \;\; 0 < |\epsilon| < 1. \tag{1.12}$$

Figure 0.5.

If there was a proper moduli scheme for elliptic curves (a concept we shall discuss further below, but for the time being can be thought of as a complete variety parameterizing isomorphism classes of elliptic curves) this family would have to have a limit at zero. The obvious way to complete this family is by taking the fiber above zero to be $y^2 = x^2(x-5)$. This is no longer an elliptic curve. However, if we delete the singular point $(0,0)$ the remaining variety is a family of algebraic groups (with the identity point always being the point at infinity) over $|\epsilon| < 1$. The fiber above zero being in fact isomorphic to $\mathbb{G}_m$. We obtain a family of elliptic curves degenerating to the multiplicative group $\mathbb{G}_m$. The theory of compactification of moduli spaces of abelian varieties is the content of [**31**].

DEFINITION 1.8. Let G, H be two algebraic groups over k. We let $\mathrm{Hom}(G,H)$ denote the homomorphisms as algebraic groups from G to H defined over any extension of k. We let $\mathrm{Hom}_k(G,H)$ stand for those defined over k.

Note that $\mathrm{Hom}_{k^s}(G,H)$ is a Γ-set. The action is given by

$$({}^{\gamma}h)(x) = \gamma(h(\gamma^{-1}x)). \tag{1.13}$$

If we write G as a quasi-projective variety in $\mathbb{P}^n_k$, then every $x \in G(K)$ can be thought of as a $n+1$ tuple $(x_0 : \cdots : x_n)$ with $x_i \in K$. Then for $\sigma \in \Gamma$ we have $\sigma x = (\sigma x_0 : \cdots : \sigma x_n)$ (there is a unique extension of the Γ-action on $G(k^s)$ to $G(\overline{k})$). However, the Galois action can be defined intrinsically:

Let X be a scheme over the field F, where F is either k^s or K. Suppose that X can be defined over k. That is, there is a scheme X_0 over k such that $X = X_0 \times_k F$ (we use often this shorthand for $X \times_{\mathrm{Spec}(k)} \mathrm{Spec}(F)$). Then for every Galois automorphism $\gamma \in \Gamma$ we can define a conjugation morphism

$$\tau_\gamma : X \longrightarrow X, \tag{1.14}$$

as follows. Let $\phi : \mathrm{Spec}(F) \longrightarrow \mathrm{Spec}(F)$ be the morphism induced from the field homomorphism $\gamma^{-1} : F \longrightarrow F$. Then τ_γ is the morphism

$$X_0 \times_k F \xrightarrow{id_{X_0} \times \phi} X_0 \times_k F \ . \tag{1.15}$$

We remark that if $X_0 = \mathrm{Spec}(k[x_1, \ldots, x_n]/(f_1, \ldots, f_m))$ then

$$X = \mathrm{Spec}(F[x_1, \ldots, x_n]/(f_1, \ldots, f_m)), \tag{1.16}$$

the isomorphism of X with $X_0 \times_k F$ is coming from the ring isomorphism

$$F[x_1, \ldots, x_n]/(f_1, \ldots, f_m) \cong k[x_1, \ldots, x_n]/(f_1, \ldots, f_m) \otimes_k F. \tag{1.17}$$

The morphism

$$X \xrightarrow{\tau_\gamma} X, \tag{1.18}$$

is induced from

(1.19)

$$k[x_1, \ldots, x_n]/(f_1, \ldots, f_m) \otimes_k F \xleftarrow{a \otimes \gamma^{-1}(r) \leftarrow a \otimes r} k[x_1, \ldots, x_n]/(f_1, \ldots, f_m) \otimes_k F \ .$$

Thus, τ_γ takes the ideal $(x_1 - \alpha_1, \ldots, x_n - \alpha_n)$ to the ideal $(x_1 - \gamma(\alpha_1), \ldots, x_n - \gamma(\alpha_n))$. That is, the effect of τ_γ on points (in the naive sense) is just

$$\tau_\gamma(\alpha_1, \ldots, \alpha_n) = (\gamma(\alpha_1), \ldots, \gamma(\alpha_n)). \tag{1.20}$$

N.B. Although τ_γ is a morphism of schemes, it is not a morphism of schemes over F.

If X_0, Y_0 are two schemes over k then Γ acts on $\mathrm{Hom}_F(X_0 \times_k F, Y_0 \times_k F)$ by the rule

$$\phi \mapsto \tau_\gamma \circ \phi \circ \tau_\gamma^{-1}. \tag{1.21}$$

In particular, getting back the case algebraic groups G, H, defined over k, one can prove

$$\mathrm{Hom}_k(G, H) = \mathrm{Hom}_{k^s}(G, H)^\Gamma. \tag{1.22}$$

This is descent argument. We refer the reader to [**78**] and [**6**] for more on Galois action and descent. Note that if H is commutative then $\mathrm{Hom}(G, H)$ is a Γ-module.

Definition 1.9. The *characters* of G are the group

$$X(G) = \mathrm{Hom}(G, \mathbb{G}_m). \tag{1.23}$$

Note that if $f \in X(G)$ then, in particular, f is a non-vanishing regular function on G. Thus, for abelian varieties, $X(G)$ is trivial (since any morphism from a projective variety to an affine one is constant). However, as mentioned above, abelian varieties may degenerate into affine groups, and even into affine groups G for which $X(G)$ determines completely the group G. These are called diagonalizable groups and are discussed extensively below.

DEFINITION 1.10. The *co-characters* (or, multiplicative one parameter subgroups) of an algebraic group G are

$$X_*(G) = \mathrm{Hom}(\mathbb{G}_m, G). \tag{1.24}$$

There is a pairing

$$X_* \times X \longrightarrow \mathrm{Hom}(\mathbb{G}_m, \mathbb{G}_m) = \mathbb{Z}. \tag{1.25}$$

EXERCISE 1.11. Prove that $X(\mathrm{GL}_2)$ is a free abelian group of rank 1 generated by the determinant character. Find $X_*(\mathrm{GL}_2)$ and the pairing $X_* \times X \longrightarrow \mathbb{Z}$ explicitely.

LEMMA 1.12. (Independence of characters) *The subset $X(G) \subset K[G]$ is linearly independent over K.*

PROOF. It is a classical theorem due to Dedekind. See [**5**, Lemma 8.1, p.111] □

DEFINITION 1.13. We say that G is *diagonalizable* if $K[G] = K[X(G)]$, where $K(G)$ is the affine coordinate ring of G and $K[X(G)]$ is the linear span over K of $X(G)$. If furthermore, $X(G)_k := \mathrm{Hom}_k(G, \mathbb{G}_m)$ spans $k[G]$, we say that G is *split over k*.

THEOREM 1.14. *Assume $Y \subset X(G)_k$ spans $k[G]$. Then:*

1. *$Y = X(G)$. In particular, $X(G)_k = X(G)$.*
2. *$k[G] = k[X(G)]$ as Hopf algebras, where the right hand side stands for the group algebra of the group $X(G)$ with co-multiplication induced from the diagonal map*

$$\Delta : X(G) \longrightarrow X(G) \times X(G), \quad \Delta(f) = (f, f), \tag{1.26}$$

and co-inverse given by

$$\widetilde{inv} : X(G) \longrightarrow X(G), \ \ \widetilde{inv}(f) = f^{-1}. \tag{1.27}$$

3. *If H is a closed subgroup of G, then H is diagonalizable and is split over k .*

PROOF.

1. We have $k[G] = k[Y] \subset k[X(G)_k] \subset k[G]$ and $X(G)_k$ is independent over k. Thus $Y = X(G)_k$. Applying that to $Y = X(G)_k \subset X(G)_K$ we get the rest.
2. The only thing to verify is that the structure on both sides agree. This is straightforward. The algebra structure is provided by the definition. The co-algebra follows from these considerations: if $\tilde{m} : k[G] \longrightarrow k[G] \otimes k[G]$ is co-multiplication, then $(\tilde{m}f)(x, y) = f(x * y)$, where $*$ is multiplication in G. If f is in $X(G)$, then $f(x * y) = f(x)f(y)$. This shows that $\tilde{m}f = f \otimes f$, i.e., the diagonal morphism on $X(G)$ induces $\tilde{m}f = f \otimes f$.

3. By definition, $K[H]$ is a homomorphic image of $K[G] = K[X(G)]$ and is thus spanned by the images of $X(G)$, which are characters on H. Thus, H is diagonalizable, $K[H] = K[X(H)]$, and $K[G] \longrightarrow K[H]$ is the group algebra epimorphism $\psi_* : K[X(G)] \longrightarrow K[X(H)]$ induced by the group homomorphism $\psi : X(G) \longrightarrow X(H)$. But then, since $X(G) = X(G)_k$, it is clear that the kernel is defined over k and $X(H) = X(H)_k$. In fact, $\mathrm{Ker}(\psi_*)$ is generated by $\{m - \mathbf{1} | m \in \mathrm{Ker}(\psi)\}$.

□

REMARK 1.15. 1. Under the assumptions of the theorem, since G is of finite type over k, $X(G)$ is finitely generated, and therefore we have a surjective group algebra homomorphism $k[\mathbb{G}_m^r] = k[\mathbb{Z}^r] \longrightarrow k[X(G)] = k[G]$ for some $r \in \mathbb{N}$. Thus, G is a closed subgroup of ${\mathbb{G}_m}^r$. One views ${\mathbb{G}_m}^r$ as embedded diagonally in GL_r, hence the name "diagonalizable". Conversely, Theorem 1.14 says that any closed subgroup of ${\mathbb{G}_m}^r$ is diagonalizable in our sense.
2. Let G be diagonalizable and $k \subset k' \subset K$. Then G is split over k' iff $X(G) = X(G)_{k'}$.

PROPOSITION 1.16. *The contravariant functor,*

$$G \mapsto X(G), \tag{1.28}$$

from the category of k-split diagonalizable groups and k-homomorphisms of algebraic groups, to the category of finitely generated $\mathbb{Z}$-modules and homomorphisms of modules is fully faithful.

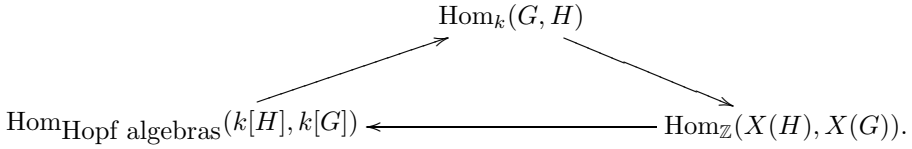
PROOF. We have a natural map

$$\mathrm{Hom}_k(G, H) \longrightarrow \mathrm{Hom}_{\mathbb{Z}}(X(H), X(G)). \tag{1.29}$$

The claim we are making is that it is an isomorphism. In particular, (exchanging k with K), we see that all the homomorphisms between such groups are defined over k.

One simply verifies that any closed circle (starting from any point) in the following triangle is the identity:

(1.30)



□

DEFINITION 1.17. An algebraic group G isomorphic over K to ${\mathbb{G}_m}^n$ is called an *n-dimensional torus*.

Let G be an n-dimensional torus, then G is a connected diagonalizable group of dimension n. As remarked before, if G is diagonalizable and $X(G) \cong \mathbb{Z}^n$ then G is a n-dimensional torus. Suppose that G is a connected diagonalizable group of dimension n. Since $\mathbb{G}_m$ is connected of dimension 1, its only connected reduced

subgroups are 1 and $\mathbb{G}_m$. (In characteristic p the assumption of reduced is important. See Example 1.23). Thus every element of $X(G)$ is either trivial or surjective. Hence, $X(G)$ is a free abelian group of rank n. We proved:

PROPOSITION 1.18. *The following are equivalent:*

1. G is a n-dimensional torus;
2. G is a connected diagonalizable group of dimension n;
3. G is diagonalizable and $X(G) \cong \mathbb{Z}^n$.

COROLLARY 1.19. *A closed connected subgroup of a torus is a torus and is a direct factor.*

PROOF. Apply Proposition 1.16 and 1.18. □

FACT 1.20. Every diagonalizable group G over k is split over k^s, thus over a finite extension. (Essentially because Γ is compact (being a profinite group) and the representation ρ associated to $X(G)$ is continuous with discrete image, hence finite; the group G is split over L, the fixed field of $\mathrm{Ker}(\rho)$, which is a finite field extension of k.)

Recall that a bilinear group homomorphism $\phi : A \times B \longrightarrow C$ is called a *perfect pairing* if

$$\forall a \in A, \phi(a,b) = 0 \implies b = 0 \tag{1.31}$$

and if

$$\forall b \in B, \phi(a,b) = 0 \implies a = 0. \tag{1.32}$$

EXAMPLE 1.21. Let A be a finite abelian group, $B = A^* = \mathrm{Hom}(A, \mathbb{C}^\times)$. Then the pairing

$$A \times B \longrightarrow \mathbb{C}^\times, \quad (a, \psi) \mapsto \psi(a). \tag{1.33}$$

is perfect.

PROPOSITION 1.22. *Let T be a torus defined over k, $\Gamma = \mathrm{Gal}(k^s/k)$.*
a) $X(T) = X(T)_{k^s}$, $X_* = X_*(T)_{k^s}$. Hence,

$$X(T)_k = X(T)^\Gamma \quad , \quad X_*(T)_k = X_*(T)^\Gamma. \tag{1.34}$$

b) The pairing

$$X_*(T) \times X(T) \longrightarrow \mathbb{Z} \tag{1.35}$$

is a perfect pairing of Γ-modules.

PROOF. By Proposition 1.16, for any k^s-split diagonalizable group G we have

$$\mathrm{Hom}(G,T) = \mathrm{Hom}_{k^s}(G,T). \tag{1.36}$$

In particular, $X_* = X_*(T)_{k^s}$, and hence $X_*(T)_k = X_*(T)^\Gamma$. We leave the proof of b) to the reader as an exercise (be sure to check also the compatibility with the action of $\gamma \in \Gamma$). □

Let G be a diagonalizable k-group (but not necessarily k-split). Since G splits over k^s, we have

$$K[G] = K \otimes_{k^s} k^s[X(G)] \quad , \quad k^s[G] = k^s[X(G)]. \tag{1.37}$$

The action of Γ on an element $\sum a_\alpha \alpha \in k^s[X(G)]$ is

$$\gamma(\sum a_\alpha \alpha) = \sum (\gamma a_\alpha)^\gamma \alpha \quad \text{for } \gamma \in \Gamma. \tag{1.38}$$

We conclude that the Γ-module $X(G)$ determines $k[G] = k^s[X(G)]$ uniquely. If H is another diagonalizable group, then $\mathrm{Hom}(G,H) = \mathrm{Hom}_{k^s}(G,H)$, and hence the following assertions are equivalent:

1. $h \in \mathrm{Hom}(G,H)$ is defined over k;
2. h is Γ-equivariant ;
3. $h^* : k^s[H] \longrightarrow k^s[G]$ is Γ-equivariant;
4. $X(h) : X(H) \longrightarrow X(G)$ is Γ-equivariant.

Let $\mathcal{A}$ be the category of diagonalizable groups over k with morphisms being k-homomorphisms of algebraic groups. Let $\mathcal{B}$ be the category of finitely generated continuous Γ-modules X, without p-torsion if char $k = p > 0$, with Γ-equivariant homomorphisms. Here continuous means that the map

$$\Gamma \times X \longrightarrow X \tag{1.39}$$

is continuous where Γ is given its profinite topology and X the discrete topology. Or still simpler, the action of Γ factors through a finite quotient (X being finitely generated).

We remark that for X an abelian group, $k[X]$ contains nilpotent elements if and only if char $k = p > 0$ and X has p-torsion. Since varieties are *reduced* by definition (i.e. the sheaf of regular functions contains no nilpotent elements), we have to exclude this case. To illustrate we give

Example 1.23. Suppose $k = \mathbb{Z}/p\mathbb{Z}$, and $X = \mathbb{Z}/p\mathbb{Z}$. We get

$$k[X] = k[y]/(y^p - 1), \tag{1.40}$$

the group scheme $\mu_{p,k}$ of p-th roots of unity over k, and $y-1$ is clearly nilpotent. In the context of group *schemes* $\mu_{p,k}$ is a connected diagonalizable subgroup scheme of $\mathbb{G}_m$, but in the context of group *varieties* it must be excluded. See Appendix A, Section 3.

Theorem 1.24. *The functor of characters*

$$X : \mathcal{A} \longrightarrow \mathcal{B}, \;\; G \mapsto X(G) \tag{1.41}$$

is an anti-equivalence of categories, between the category $\mathcal{A}$ of diagonalizable groups and the category $\mathcal{B}$ of continuous Γ-modules with no p-torsion if char$(k) = p$.

Proof. By proposition 1.16 and the remarks just made, the only thing we need to show is essential surjectivity, that is, every object in $\mathcal{B}$ is the character module of some object in $\mathcal{A}$.

Let M be a finitely generated Γ-module. Let G be the associated affine group with affine coordinate ring $k^s[M]$, an algebra of finite type over k^s. If $m \in M$, then m defines a character of G by

$$k^s[x, x^{-1}] \longrightarrow k^s[G] = k^s[M], \tag{1.42}$$

$$x \mapsto m. \tag{1.43}$$

(m being invertible in $k^s[M]$), and one easily verifies that in fact $M = X(G)$ (because $M \subset X(G)$ and it spans $k^s[G]$ as in Theorem 1.14). Hence, G is diagonalizable.

Consider now $(k^s[M])^\Gamma$. It satisfies

$$k^s \otimes (k^s[M])^\Gamma = k^s[M], \tag{1.44}$$

and hence gives G a k-structure. The equality (1.44) is essentially the assertion that if Γ acts continously semi-linearly, i.e.

$$\gamma(ax) = \gamma(a)\gamma(x) \quad (a \in k_s, x \in k^s[M]), \tag{1.45}$$

on a k^s-vector space V, there exists a k-basis to V. This in turn follows from

$$H^1(\Gamma, \mathrm{GL}_n(k^s)) = 1. \tag{1.46}$$

One can also prove it directly (see [**5**, Section AG 14.2]). □

REMARK 1.25. We have seen that diagonalizable groups are completely determined by their character module considered as a Γ-module. In particular, since $\mathrm{GL}_n(\mathbb{Z})$ is the automorphism group of $\mathbb{G}_m^n$, one sees that the n-dimensional tori are equivalent to Galois representations $\rho : G \longrightarrow \mathrm{GL}_n(\mathbb{Z})$ up to conjugacy. All n-dimensional tori are forms of $\mathbb{G}_m^n$ (that is, they become isomorphic to $\mathbb{G}_m^n$ over K), and in fact $H^1(\Gamma, \mathrm{Aut}(\mathbb{G}_m^n))$ – the pointed set classifying forms of $\mathbb{G}_m^n$– is indeed naturally identified with the set of Galois representations ρ modulo conjugacy.

EXERCISE 1.26. 1. Classify n -dimensional tori over the field $\mathbb{Z}/p\mathbb{Z}$. Suggestion: Do first the one and two dimensional cases.
2. If you know some class field theory, classify one dimensional tori over $\mathbb{Q}$.

REMARK 1.27. One often writes $k^s[M]$ in the form

$$k^s[X^\alpha : \alpha \in M]/(X^{\alpha+\beta} - X^\alpha X^\beta, X^0 - 1). \tag{1.47}$$

We finish this section by giving some general methods to construct tori.

EXAMPLE 1.28. *Restriction of scalars.* Let L/k be a separable field extension of degree d. We assume $L \subset k^s$, where k^s is our fixed separable closure. Let $\{\sigma_1, ..., \sigma_d\}$ be the embeddings of L into k^s. This a Γ-set:

(1.48)
$$\begin{array}{ccc} L & \xrightarrow{\sigma_i} & k^s \\ & \searrow^{\tau\circ\sigma_i} & \downarrow{\scriptstyle \tau} \\ & & k^s \end{array} .$$

Let us write

$$\tau \circ \sigma_i = \sigma_{\tau(i)}. \tag{1.49}$$

Consider the continuous Γ-module

$$M = \bigoplus_{i=1}^{d} \sigma_i \mathbb{Z}. \tag{1.50}$$

Let G be the unique, up to isomorphism, algebraic torus with $X(G) = M$. Note that G is a d-dimensional torus. It is denoted $\mathbf{Res}_{L/k}\mathbb{G}_m$.

To generalize our construction, consider an algebraic variety X/L. Define $\mathbf{Res}_{L/k}X$ as a functor

$$\mathbf{Res}_{L/k}X:\ k-\mathbf{alg}\longrightarrow \mathbf{Sets},\ \ \mathbf{Res}_{L/k}X(B)=X(B\otimes_k L). \tag{1.51}$$

One can show that there exists an algebraic variety over k, unique up to k isomorphism, such that $\mathbf{Res}_{L/k}X$ is its functor of points. This variety is denoted by $\mathbf{Res}_{L/k}X$ as well.

EXERCISE 1.29. Let L/k be a quadratic extension of fields. Write $\mathbf{Res}_{L/k}\mathbb{G}_m$ explicitly as an affine variety over k. Write also the co-multiplication and co-inverse morphisms.

To tie together the functorial and the lattice approach, note that a k^s-point of G amounts to a homomorphism $M\longrightarrow k^{s\times}$, or, with the choice of the standard basis on M, simply to a d-tuple, $(x_1,\dots,x_d)$, of elements of $k^{s\times}$. Therefore, for every field extension k' of k in k^s, we have

$$G(k')=\{(x_1,\dots,x_d):\tau(x_i)=x_{\tau(i)},\ \forall\tau\in Gal(k^s/k')\}. \tag{1.52}$$

In particular, $G(k)$ is naturally isomorphic to $L^\times$, and in fact for every separable field extension k' of k we find a natural identification of $G(k')$ with $(k'\otimes L)^\times$. That shows that G *is* $\mathbf{Res}_{L/k}\mathbb{G}_m$.

EXAMPLE 1.30. *Tensor and Hom constructions.* Let T be a d-dimensional torus over k and let $\mathcal{O}$ be the ring of integers of an algebraic number field. Let us assume that $\mathcal{O}$ is contained in the endomorphism ring of T. We remark that such examples exist in abundance, because the endomorphism ring of T over k^s is $M_d(\mathbb{Z})$.

Given a projective $\mathcal{O}$-module M of rank r, we define two new tori over k, both of dimension rd. They are denoted $T\otimes_{\mathcal{O}}M$ and $\mathrm{Hom}_{\mathcal{O}}(M,T)$. As functor of points they are defined by the following formulae: For every k-algebra B

$$(T\otimes_{\mathcal{O}}M)(B)=T(B)\otimes_{\mathcal{O}}M, \tag{1.53}$$

$$(\mathrm{Hom}_{\mathcal{O}}(M,T))=\mathrm{Hom}_{\mathcal{O}}(M,T(B)). \tag{1.54}$$

We leave it to reader to verify that these are indeed the functor of points of tori over k. We remark that the character modules are given by

$$X(T\otimes_{\mathcal{O}}M)=X(T)\otimes_{\mathcal{O}}M,\ \ X(\mathrm{Hom}_{\mathcal{O}}(M,T))=\mathrm{Hom}_{\mathcal{O}}(M,X(T)). \tag{1.55}$$

EXAMPLE 1.31. *Serre tori.* Let K be a CM field of degree $2g$ with its totally real field K^+. We remind the reader that by definition K is a totally imaginary quadratic extension of a totally real field. Let $G=\mathbf{Res}_{K/\mathbb{Q}}\mathbb{G}_m$ as in Example 1.28. The character module of G can be explicitly written as

$$X(G)=\left\{\sum n_\sigma\cdot\sigma:\sigma\in\mathrm{Hom}(K,\mathbb{Q}^s),n_\sigma\in\mathbb{Z}\right\}. \tag{1.56}$$

Let τ be complex conjugation. Then τ acts on $X(G)$. Consider the Γ-module N given by

$$N=\left\{\sum n_\sigma\cdot\sigma:n_\sigma+n_{\sigma\tau}=n_\rho+n_{\rho\tau},\ \forall\sigma,\rho\right\}. \tag{1.57}$$

Note that N is of rank $g+1$. It defines a $g+1$ dimensional torus (called a *Serre torus*) over $\mathbb{Q}$.

2. Moduli Problems

In this section we discuss moduli problems and moduli spaces. We mainly focus on ideas rather than on the solutions because the solutions – the moduli spaces themselves – are studied closely in later chapters. Some of the paragraphs in this section will be better understood only after reading on the book.

The notion of a moduli space is rather a set of examples or problems than a well defined notion. Usually, by a moduli problem one means a certain kind of classification problem. Large part of mathematics is devoted, or motivated by, classification problems: classifying topological spaces up to homotopy; classifying all the finite simple groups; classifying Hilbert spaces up to isomorphism etc. Usually a distinguishing feature of the classification problems called moduli problems is that they appear in the context of geometry, and that one considers families of objects as well.

In the context of algebraic geometry one may choose the most general approach and define a moduli problem as follows: Fix a base scheme S. A moduli problem over S is a contravariant functor

$$\Phi : \mathbf{Sch}_S \longrightarrow \mathbf{Sets} \tag{2.1}$$

from the category of schemes over S to the category of sets. Assume that this functor is *representable*, i.e., that there exists a scheme $T \longrightarrow S$ such that the functor of points of T, h_T, is naturally equivalent to Φ. That is, that there exists a natural isomorphism

$$h_T(R) := \mathrm{Mor}_S(R, T) \cong \Phi(R), \tag{2.2}$$

where R varies over S-schemes. One may then call T a *fine moduli scheme*. If T is equi dimensional over S (and nice enough such that this makes sense), then one may call $\dim_S(T)$ the number of moduli of the moduli problem Φ.

It seems that the first moduli problem, as such, was described and solved by B. Riemann. It is still a problem that motivates much research in algebraic geometry. In this case one fixes an integer $g \geq 2$ and takes Φ to be the functor associating to a scheme T the isomorphism classes of smooth curves $\mathcal{C} \longrightarrow T$ of genus g (here $S = \mathrm{Spec}(\mathbb{Z})$ and hence omitted). Riemann made the statement that this problem has $3g - 3$ moduli. Local deformation theory as developed by Kodaira-Spencer, [**63**], i.e. the theory of infinitesimal deformations (and in particular deformations over the base $k[\epsilon]$, $\epsilon^2 = 0$, k a field, called also first order deformations), dictates that the $k[\epsilon]$ deformations of a curve C are parameterized by $H^1(C, T_C)$, where T_C designates the tangent sheaf of C. By the Riemann-Roch theorem and Serre's duality

$$\dim\ H^1(C, T_C) = \dim\ H^0(C, \mathcal{O}(2K_C)) = 3g - 3. \tag{2.3}$$

See Section 6 and Chapter 3, Section 5 for more examples.

There does not exists a fine moduli scheme for this problem. It is a general philosophy that the existence of a fine moduli space that represents a functor of

isomorphism classes of certain objects, [1] usually rules out the possibility that the objects being parameterized have automorphisms, as some curves do. However, there exists a coarse moduli space for this problem.

A *coarse moduli scheme* is a scheme $T \longrightarrow S$ such that (i) there exists a morphism of functors $\Phi \longrightarrow h_T$; (ii) for every algebraically closed field k, with $\mathrm{Spec}(k) \longrightarrow S$, the morphism of functors gives a bijection $T(\mathrm{Spec}(k)) \cong \Phi(\mathrm{Spec}(k))$ and moreover (iii) every morphism $\Phi \longrightarrow h_X$ from Φ to the functor of points of a S-scheme X, factor through $\Phi \longrightarrow h_T$ for a unique morphism $T \longrightarrow X$.

Coarse moduli schemes may exist when fine moduli schemes do not. For example, there exists a coarse moduli scheme for the functor Φ of isomorphism classes of curves of genus g (still $g \geq 2$). We shall denote it by $\mathcal{M}_g$.

Often the existence and the "justification" for the existence of coarse moduli spaces is the following. One adds structure to the original problem, i.e., considers a functor Φ' with a natural ("forgetful") morphism $\Phi' \longrightarrow \Phi$, such that in fact Φ is obtained from Φ' by a "nice" equivalence relation $\sim$. For example, in the case of curves, we may take Φ' to be the moduli problem of isomorphism classes of m-pointed smooth curves. That is, a curve together with m-distinct points on it . When m is large enough, that is $m \geq N(g)$ where $N(g)$ depends only on the genus, this moduli problem is *rigid*. That is, an automorphism of a curve of genus g fixing the m marked points is necessarily the identity. The functor Φ' is representable; there exists a fine moduli space for it, say $\mathcal{M}_{g,m}$. The space $\mathcal{M}_g$ is then a quotient of $\mathcal{M}_{g,m}$ by an equivalence relation coming from forgetting the marked points.

It follows tautologically(!) that if there exists a fine moduli scheme T for a moduli problem Φ, then there exists a universal object $\mathcal{U} \longrightarrow T$. That is there exists an element $\mathcal{U} \in \Phi(T)$ with the following property: Let R be a S-scheme and $E_R : \Phi(R) \longrightarrow T(R)$ be the given bijection. Let $x \in \Phi(R)$ and $E_R(x) : R \longrightarrow T$ the corresponding morphism. By contravariance we get an object $E_R(x)(\mathcal{U}) \in \Phi(R)$. It is equal to x. In the case of m-pointed curves, this boils down to saying that there exists a universal m-pointed curve $\mathcal{U} \longrightarrow \mathcal{M}_{g,m}$ such that for every scheme R the isomorphism classes of m-pointed curves over R are precisely $\{f^*(\mathcal{U}) : f \in T(R)\}$ (remember that a point $f \in T(R)$ is, by definition, a morphism $R \longrightarrow T$).

The moduli problems we will be interested in are those of classifying abelian varieties with certain extra structure. This extra structure is meant to rigidify the moduli problem. Some moduli schemes of elliptic curves are discussed in Section 3, and of abelian varieties in Section 2.2. We do not presume to actually prove the existence of such moduli schemes. What we shall actually show is that for some moduli problems (e.g. elliptic curves and abelian varieties with real multiplication) there are certain quotients of a power of the complex upper half plane $\mathcal{H}$ that are "likely" to be coarse moduli schemes. In fact they are, but we shall only demonstrate a natural bijection between the points of such quotients and the value of the respective moduli problem on the scheme $\mathrm{Spec}(\mathbb{C})$. This falls short of actually proving they are coarse moduli schemes.

We want to draw the reader's attention to a phenomenon she encountered before (or so we expect): There is no universal elliptic curve.

[1] Without the proviso on the functor representing isomorphism classes it is easy to give examples where automorphisms form no obstruction. Grassmannians and formal group laws are such examples

Recall that the coarse moduli scheme of elliptic curves is given by the j-line; two elliptic curves are isomorphic if and only if they have the same j invariant. There doesn't exist an elliptic curve over the j-line such that its fibre over every j_0 is the elliptic curve with invariant j_0. This has to do, of course, with the j-line not being quite a fine moduli scheme. There is, however, such an elliptic curve over the j-line with the points 0 and 1728 removed: E.g., $y^2 + xy = x^3 - \frac{36}{j-1728}x - \frac{1}{j-1728}$. If we ask for the moduli of elliptic curves with a 7 torsion point (this is a rigid problem) then there exists a fine moduli scheme and thus a universal elliptic curve over it. This moduli scheme is an open sub-variety of $\mathbb{P}^1$. Given $t \in \mathbb{A}^1$ associate to it the elliptic curve $y^2 + (1 + t - t^2)xy + (t^2 - t^3)y = x^3 + (t^2 - t^3)x^2$. It has discriminant $t^7(t-1)^7(t^3-8t^2+5t+1)$. The moduli scheme is then the complement of the zero locus of the discriminant. The 7 torsion point is $(0, 0)$. We have taken those examples from Silverman [**107**, Chapter III.1, Appendix C.13].

For elliptic curves, if one takes as extra structure, say, an isomorphism of the n-torsion points ($n \geq 3$) with $(\mathbb{Z}/n\mathbb{Z})^2$ then one gets a rigid moduli problem representable by a fine moduli scheme over $\mathbb{Z}[\zeta_n]$. Although the same level structure is rigid (that is, objects do not have any automorphism except the identity) for abelian varieties one needs to put some more structure to obtain nice moduli schemes. The reason has to do with the way one construct these moduli schemes.

The usual procedure is to realize all the objects one parameterizes as sub varieties of a projective space of a fixed dimension, having special properties. Thus they are typically parameterized a sub variety of an appropriate Hilbert scheme. For curves of genus greater then 1, the divisor $3K$ (K the canonical divisor) is very ample, and thus all curves of genus g appear as sub varieties of $\mathbb{P}^{5g-6}$. See [**20**, Section 1]. For elliptic curves, $3 \cdot 0_E$ is a very ample divisor, and all elliptic curves appear as sub varieties of $\mathbb{P}^2$ (Weierstrass equation). One proceeds to construct the moduli scheme as a sub-variety of a quotient of the appropriate Hilbert scheme.

Giving a projective embedding into a fixed projective space requires a choice of a very ample divisor and a choice of basis for its global sections. All these are extra data and one attempt to discard it. The choice of basis is corrected by dividing the Hilbert scheme by the action of PGL_n. The choice of ample divisor is corrected by considering divisors up to algebraic equivalence. For elliptic curves, the last correction is not needed because there is a canonical choice of an ample divisor, viz., 0_E. However, for abelian varieties there is no such choice and one needs to choose an ample divisor up to algebraic equivalence. Since two divisors on an abelian variety are algebraically equivalent if and only if they define the same map from the abelian variety to its dual, one finds the usual definition of a polarization as a map from the abelian variety to its dual induced by an ample line bundle (see Definitions 1.14 and 6.27). We note that while for elliptic curves the Néron-Severi group – the group of divisors modulo algebraic equivalence – is just $\mathbb{Z}$, for an abelian variety, though it is still a torsion free abelian group, it may happen that it is of rank greater than 1 and thus there is no natural choice of polarization.

Thus, a typical moduli problem for abelian varieties is that of parameterizing triples (A, λ, γ) where A is an abelian variety of a given dimension (otherwise the moduli scheme would be a disjoint union according to dimension), a polarization $\lambda : A \longrightarrow A^t$ (that is: a map from A to its dual abelian variety that comes from an ample line bundle) which should be thought of as an equivalence class of projective

embeddings, and a level structure γ (Typically a choice of a point, or several points on the abelian variety that rigidifies the moduli problem).

3. Moduli of Elliptic Curves

In this section, as the title suggests, we examine the moduli schemes corresponding to pairs composed of an elliptic curve and an associated level structure. We assume the reader is familiar with the basic theory of elliptic curves.

Every abelian variety of dimension 1 is an elliptic curve. By an elliptic curve E over a ring R, we mean a group scheme E over $\mathrm{Spec}(R)$, of relative dimension one, which is proper with geometrically connected fibers. Equivalently, locally on $\mathrm{Spec}(R)$, the scheme E can be given a Weierstrass equation $y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$ with variables x, y.

DEFINITION 3.1. Let R be a ring in which n is invertible and assume that R is a $\mathbb{Z}[\zeta_n]$-module ($\zeta_n = e^{2\pi i/n}$).

1. A *full symplectic level-n-structure* (or a $\Gamma(n)$*-structure*) on an elliptic curve E over R is an isomorphism of constant group schemes over R

$$\alpha : (\mathbb{Z}/n\mathbb{Z})^2 \longrightarrow E[n]. \tag{3.1}$$

   We further require that under α the Weil pairing on $E[n]$ induces the symplectic pairing on $(\mathbb{Z}/n\mathbb{Z})^2$ determined by $< (1,0),(0,1) >= \zeta_n$.
2. *A* $\Gamma_1(n)$*-level structure* is a choice of a point $P \in E[n](R)$ of order n.
3. A $\Gamma_0(n)$*-level structure* is a choice of cyclic subgroup $H \subset E(n)$ of order n that is defined over R.

REMARK 3.2. 1. The phrase "constant group schemes over R" amounts to saying that all the n-torsion points of E are defined over R.
2. By a symplectic pairing of $(\mathbb{Z}/n\mathbb{Z})^2$ we mean a map

$$\phi : (\mathbb{Z}/n\mathbb{Z})^2 \times (\mathbb{Z}/n\mathbb{Z})^2 \longrightarrow \mu_n(R), \tag{3.2}$$

   such that ϕ is bilinear and antisymmetric.

We consider the following functors:

$$M(n) : R \mapsto M(n)(R) := \{\text{isomorphism classes over } R \text{ of } (E,\alpha)\} \tag{3.3}$$

$$M_1(n) : R \mapsto M_1(n)(R) := \{\text{isomorphism classes over } R \text{ of } (E,P)\} \tag{3.4}$$

$$M_0(n) : R \mapsto M_0(n)(R) := \{\text{isomorphism classes over } R \text{ of } (E,H)\} \tag{3.5}$$

To clarify, two pairs (E_1,α_1) and (E_2,α_2) in $M(n)(R)$ are isomorphic over R if there exists an isomorphism $f : E_1 \longrightarrow E_2$, defined over R, such that $f \circ \alpha_1 = \alpha_2$. Similarly for $\Gamma_1(n)$ and $\Gamma_0(n)$ level structures. Thus, we always have $(E,P) \cong (E,h(P))$, for $h \in \mathrm{Aut}_R(E)$. For most elliptic curves h could only be ± 1, but for some curves there are more possibilities.

We have the following natural transformations of functors:

$$M(n) \longrightarrow M_1(n), \tag{3.6}$$

given by

$$(E,\alpha) \mapsto (E,\alpha(1,0)), \tag{3.7}$$

and

$$M_1(n) \longrightarrow M_0(n), \tag{3.8}$$

given by

$$(E, P) \mapsto (E, < P >), \tag{3.9}$$

where $< P >$ means the cyclic group generated by P.

Consider now elliptic curves over $\mathbb{C}$. Let $\mathcal{H}$ be the complex upper half plane

$$\mathcal{H} = \{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}. \tag{3.10}$$

Let $\tau \in \mathcal{H}$, and let $\mathcal{L}_\tau$ be the lattice corresponding to τ, i.e.,

$$\mathbb{Z} + \mathbb{Z} \cdot \tau. \tag{3.11}$$

Let $\mathcal{E}_\tau$ be the elliptic curve

$$\mathbb{C}/\mathcal{L}_\tau. \tag{3.12}$$

Later, in Example 6.39, we shall see that every elliptic curve is isomorphic to $\mathcal{E}_\tau$ for a suitable $\tau \in \mathcal{H}$, and we shall prove that under the Weil pairing

$$<,>: \mathcal{E}_\tau[n] \times \mathcal{E}_\tau[n] \longrightarrow \mu_n, \tag{3.13}$$

we have

$$< \frac{1}{n}, \frac{\tau}{n} >= \exp(2\pi i/n) = \zeta_n. \tag{3.14}$$

Granted that, we get a symplectic level-n-structure

$$\alpha : (\mathbb{Z}/n\mathbb{Z})^2 \longrightarrow \mathcal{E}_\tau[n], \tag{3.15}$$

by setting $\alpha(1,0) = \frac{1}{n}$ and $\alpha(0,1) = \frac{\tau}{n}$. We also get a:

- $\Gamma_1(n)$ level structure by taking P_τ to be $\alpha(1,0)$ on $\mathcal{E}_\tau$.
- $\Gamma_0(n)$ level structure by taking H_τ to be the subgroup $\left\{\frac{a}{n} | a = 0, \ldots, n-1\right\}$ on $\mathcal{E}_\tau$.

Let $\Gamma(n), \Gamma_1(n), \Gamma_0(n)$ be the following congruence subgroups of $\mathrm{SL}_2(\mathbb{Z})$:

$$\Gamma(n) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \bmod n \right\}, \tag{3.16}$$

$$\Gamma_1(n) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} \bmod n \right\}, \tag{3.17}$$

$$\Gamma_0(n) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \bmod n \right\}. \tag{3.18}$$

Let $\mu = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$, $\tau' = \mu\tau$. A short calculation yields that τ and τ' define isomorphic couples in the following settings:

$$(E_\tau, \alpha_\tau) \cong (E'_\tau, \alpha'_\tau) \iff \mu \in \Gamma(n), \tag{3.19}$$

$$(E_\tau, P_\tau) \cong (E'_\tau, P'_\tau) \iff \mu \in \Gamma_1(n), \tag{3.20}$$

$$(E_\tau, H_\tau) \cong (E'_\tau, H'_\tau) \iff \mu \in \Gamma_0(n). \tag{3.21}$$

Moreover, we have the following diagram:

$$\begin{array}{ccc} \Gamma(N)\backslash\mathcal{H} & \xrightarrow[\alpha(1,0)=\frac{1}{n},\ \alpha(0,1)=\frac{\tau}{n}]{\tau\mapsto\mathbb{C}/(\mathbb{Z}+\mathbb{Z}\tau)} & M(N)(\mathbb{C}) \\ \downarrow & & \downarrow \\ \Gamma_1(N)\backslash\mathcal{H} & \xrightarrow[P=\frac{1}{n}]{\tau\mapsto\mathbb{C}/(\mathbb{Z}+\mathbb{Z}\tau)} & M_1(N)(\mathbb{C}) \\ \downarrow & & \downarrow \\ \Gamma_0(N)\backslash\mathcal{H} & \xrightarrow[H=<\frac{1}{n}>]{\tau\mapsto\mathbb{C}/(\mathbb{Z}+\mathbb{Z}\tau)} & M_0(N)(\mathbb{C}) \end{array} \tag{3.22}$$

EXERCISE 3.3. Take a point τ for which one of the coverings is ramified. Express that in terms of automorphisms of the elliptic curve $\mathcal{E}_\tau$.

We recall that the open Riemann surfaces $\Gamma(N)\backslash\mathcal{H}, \Gamma_1(N)\backslash\mathcal{H}$ and $\Gamma_0(N)\backslash\mathcal{H}$ denoted customarily by $Y(N)(\mathbb{C}), Y_1(N)(\mathbb{C})$ and $Y_0(N)(\mathbb{C})$ respectively can be compactified canonically by adding the orbits of $\mathbb{P}^1(\mathbb{Q})$ under the group $\Gamma(N), \Gamma_1(N)$ and $\Gamma_0(n)$ respectively. The resulting curves are denoted accordingly $X(N)(\mathbb{C}), X_1(N)(\mathbb{C})$ and $X_0(N)(\mathbb{C})$.

For sake of completeness we give information of indices and genus. ****

4. Modular Forms

In this section, we present modular forms as sections of line bundles. Though our presentation is quite sell-contained, we assume that the reader is familiar with the usual definition of modular forms, and our intention is to present a re-interpretation and some interesting phenomena.

Let Γ be a discrete subgroup of $\mathrm{SL}_2(\mathbb{R})$.

DEFINITION 4.1. A holomorphic function $j : \Gamma \times \mathcal{H} \longrightarrow \mathbb{C}^\times$ is called a *factor of automorphy* if for all $\mu_1, \mu_2 \in \Gamma$ and $\tau \in \mathcal{H}$:

$$j(\mu_1\mu_2, \tau) = j(\mu_1, \mu_2\tau)j(\mu_2, \tau). \tag{4.1}$$

One calls (4.1) the *cocycle relation*. Indeed, if we view Γ as acting on the multiplicative abelian group $\mathcal{O}^\times_{\mathcal{H}}$ of non-vanishing holomorphic functions on $\mathcal{H}$ by the rule $(f^\gamma)(\tau) = f(\gamma\tau)$, then the association

$$\Gamma \longrightarrow \mathcal{O}^\times_{\mathcal{H}}, \quad \mu \mapsto j(\mu, \cdot) \tag{4.2}$$

is a 1-cocycle in $Z^1(\Gamma, \mathcal{O}^\times_{\mathcal{H}})$.

EXAMPLE 4.2. Let k be an integer. Let

$$j_k(\mu, \tau) : \Gamma \times \mathcal{H} \longrightarrow \mathbb{C}, \quad j_k(\mu, \tau) = (c\tau + d)^k, \tag{4.3}$$

where $\mu = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma, \tau \in \mathcal{H}$. Then j_k is a factor of automorphy.

The following diagram demonstrates the compatibility of the action of Γ on $\mathcal{H} \times \mathbb{C}$:

Figure 1.

DEFINITION 4.3. For $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ torsion free (i.e. Γ contains no elliptic elements), we define the Hodge bundle $\mathbb{E}^k$ as the quotient:

$$\mathcal{H} \times \mathbb{C}/\Gamma, \tag{4.4}$$

with the action of Γ given by $(\tau, \alpha) \sim_\mu (\mu\tau, j_k(\mu, \tau)\alpha)$

Some remarks are in order: First, recall that a group G acts *properly and discontinuously* on a topological space X, if for any pair of compact subsets $K_1, K_2 \subset X$ the set

$$\{g \in G | gK_1 \cap K_2 \neq \emptyset\} \tag{4.5}$$

is finite. Also, the action of G on X is said to be *free*, if $gx = x$ for some $x \in X$ and $g \in G$ implies g is the identity. So, using that Γ acts freely and properly discontinuously on $\mathcal{H} \times \mathbb{C}$, we get that the Hodge bundle $\mathbb{E}^k$ is well-defined as a complex manifold. And since Γ is torsion free, $\mathbb{E}^k$ is a line bundle over $\Gamma\backslash\mathcal{H}$.

Secondly, we note that $\mathbb{E}^{\otimes kk'} \cong (\mathbb{E}^{\otimes k})^{\otimes k'}$, because, in general, if L_i, $i = 1, 2$, are two line bundles defined by factors of automorphy h_i, then $L_1 \otimes L_2$ is defined by the factor of automorphy $h_1 h_2$.

DEFINITION 4.4. A *modular form* of weight k and level Γ is a global section of the line bundle $\mathbb{E}^k$ over $\Gamma\backslash\mathcal{H}$, i.e., an element of $H^0(\Gamma\backslash\mathcal{H}, \mathbb{E}^k)$. Equivalently, a function

$$f : \mathcal{H} \longrightarrow \mathbb{C}, \tag{4.6}$$

such that

$$f(\mu\tau) = j_k(\mu, \tau) f(\tau). \tag{4.7}$$

We denote the vector space of modular forms of weight k and level Γ by $\mathcal{F}(\mathbb{C}, k, \Gamma)$.

REMARK 4.5. Some care has to be taken here. A modular form in the sense above is meromorphic at the cusps. The Hodge bundle extends to the compactified curve $\Gamma\backslash\mathcal{H}^*$ and a modular form of weight k and level Γ is usually taken to be a section of $\mathbb{E}^k$ over $\Gamma\backslash\mathcal{H}^*$. I.e., one adds the requirement of being holomorphic at the cusps. To distinguish we shall say "holomorphic modular forms" and use the notation $\mathcal{M}(\mathbb{C}, k, \Gamma)$.

Starting with the property (4.7), one may define modular forms of weight k and level Γ for Γ not necessarily torsion free, e.g., for $\mathrm{SL}_2(\mathbb{Z})$, as holomorphic functions

$$f : \mathcal{H} \longrightarrow \mathbb{C}, \tag{4.8}$$

satisfying

$$f(\mu\tau) = j_k(\mu,\tau)f(\tau). \tag{4.9}$$

Again, the usual terminology requires holomorphity at infinity as well. In the case where Γ is not torsion free, these are the global sections of a sheaf on $\Gamma\backslash\mathcal{H}$ which is generally *not* invertible. The problem is with the ramification of the map $\mathcal{H} \longrightarrow \Gamma\backslash\mathcal{H}$ that obstructs the descent of the Hodge bundle to the quotient. To see that indeed there *is* an obstruction take the case $\Gamma = \mathrm{SL}_2(\mathbb{Z})$ and note that

$$\begin{aligned}
&\text{from } \mu_0 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \text{ we get} \quad && f(\mu_0\tau) = f(\tau) = (-1)^k f(\tau) \\
&\text{from } \mu_1 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \text{ we get} \quad && f(\mu_1 i) = f(i) = i^k f(i) \\
&\text{from } \mu_2 = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix} \text{ we get} \quad && f(\mu_2\omega) = f(\omega) = \omega^k f(\omega)
\end{aligned}$$

($\omega = \exp(2\pi i/3)$) and hence every section of the sheaf of modular forms of weight k and level $SL_2(\mathbb{Z})$ vanishes: (i) identically, if $2 \nmid k$; (ii) at i, if $4 \nmid k$; (iii) at ω, if $3 \nmid k$. On the other hand, it is well known that $\mathrm{SL}_2(\mathbb{Z})\backslash\mathcal{H} \cong \mathbb{C}$. The isomorphism is given in fact by the j-invariant. Since $\mathbb{C}$ is simply connected, any invertible sheaf over it (or equivalently, any line bundle) is trivial. But triviality of the line bundle implies the existence of a non-vanishing section! That is impossible, since $\mu_1, \mu_2 \in \mathrm{SL}_2(\mathbb{Z})$. See also the digression on factors of automorphy in Section 6.1

Exercise 4.6. Show that $\mathbb{E}^{12}$ descends to a line bundle on $\mathrm{SL}_2(\mathbb{Z})\backslash\mathcal{H}$. How does that fit with the existence of the cusp form Δ?

Interesting references in this context are [**77**] and [**79**].

We may interpret the Hodge bundle $\mathbb{E}^1$ in the following way: Given $\tau \in \mathcal{H}$, $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$, we have an isomorphism of elliptic curves (induced by multiplication):

$$\mathcal{E}_\tau = \mathbb{C}/\mathbb{Z} + \mathbb{Z}\tau \xleftarrow[\times(c\tau+d)=j(\gamma,\tau)]{\sim} \mathcal{E}_{\gamma\tau} = \mathbb{C}/\mathbb{Z} + \mathbb{Z}\Big(\tfrac{a\tau+b}{c\tau+d}\Big), \tag{4.10}$$

which induces a map between the tangent spaces

$$\mathbb{C} = \mathfrak{t}_{\mathcal{E}_{\tau,0}} \xleftarrow[\times c\tau+d]{} \mathfrak{t}_{\mathcal{E}_{\mu\tau,0}} = \mathbb{C}, \tag{4.11}$$

and therefore, a map between the cotangent spaces

$$\mathfrak{t}^*_{\mathcal{E}_{\tau,0}} \xrightarrow[\times c\tau+d]{\sim} \mathfrak{t}^*_{\mathcal{E}_{\mu\tau,0}}. \tag{4.12}$$

Consider the projection map

$$\mathcal{E}_{univ,\Gamma} := \Gamma\backslash\tilde{\mathcal{E}}_{univ,\Gamma} \xrightarrow{\ \pi\ } \Gamma\backslash\mathcal{H}, \tag{4.13}$$

where $\tilde{\mathcal{E}}_{univ,\Gamma} = \{(\tau, x) : x \in \mathbb{C}/\mathbb{Z} + \mathbb{Z}\tau\}$, the action of Γ being given by $(\tau, x) \sim (\gamma\tau, j(\gamma,\tau)^{-1}x)$. We see that the family of elliptic curves $\tilde{\mathcal{E}}_{univ,\Gamma}$ over $\mathcal{H}$ descend to a family of elliptic curves $\mathcal{E}_{univ,\Gamma}$ over $\Gamma\backslash\mathcal{H}$ (which is the "universal family with level Γ") and that the relative tangent space of $\mathcal{E}_{univ,\Gamma}$ over $\Gamma\backslash\mathcal{H}$ is defined by the factor of automorphy $j(\gamma,\tau)^{-1}$.

Conclusion: Let Γ be torsion free. A modular form of weight k and level Γ is a section of $e^*(\mathfrak{t}^*_{\mathcal{E}_{univ,\Gamma}}{}^{\otimes k})$, where $e : \Gamma\backslash\mathcal{H} \longrightarrow \mathcal{E}_{univ,\Gamma}$ is the identity section induced from $\tau \mapsto (\tau, 0)$.

REMARK 4.7. The reader may notice how the fact that there exists a universal family is connected to the fact that the map $\mathcal{H} \longrightarrow \mathcal{H}/\Gamma$ is unramified and to the fact that the automorphism group of an elliptic group with "Γ structure" is trivial. See Exercise 3.3.

Let f be a modular form of weight k and level Γ. Taking the last perspective, we have therefore a map:

$$(E, \Gamma)_R \overset{f}{\longmapsto} f(E, \Gamma) \in H^0(E, \underline{\omega}^{\otimes k}). \tag{4.14}$$

Here E is an elliptic curve over a $\mathbb{C}$-algebra R and $\underline{\omega}$ is the sheaf over $\mathrm{Spec}(R)$ of relative holomorphic forms on E. That is, $\underline{\omega} = \pi_*\Omega^1_{E/R}$. It is a projective rank 1 module over R. Every such holomorphic differential form is translation invariant and therefore there is a canonical identification of $\underline{\omega}$ with the relative cotangent sheaf at zero $\mathfrak{t}^*_{E/R,0}$.

Suppose we choose ω_0 to be some non-vanishing differential on E, and consider:

$$(E, \Gamma, \omega_0) \overset{f}{\longmapsto} \frac{f(E, \Gamma)}{\omega_0^k} \in R. \tag{4.15}$$

Replacing ω_0 with $\lambda\omega_0$, for $\lambda \in R^\times$, we get:

$$(E, \Gamma, \lambda\omega_0) \overset{f}{\longmapsto} \frac{f(E, \Gamma)}{\lambda\omega_0^k} = \lambda^{-k} f(E, \Gamma, \lambda\omega_0), \tag{4.16}$$

so we may think of a modular form f (with level Γ, weight k) as a rule

$$(E, \Gamma, \omega_0)_R \overset{f}{\longmapsto} f(E, \Gamma, \omega_0) \in R \tag{4.17}$$

such that $f(E, \Gamma, \lambda\omega_0) = \lambda^{-k} f(E, \Gamma, \lambda\omega_0)$ and f depends only on the isomorphism class of (E, Γ, ω_0) and commutes with base change. The point of this gymnastics is that such a definition makes sense if we replace $\mathbb{C}$ by any base scheme B and allow R to be any B-algebra! This insight is due to Katz.

REMARK 4.8. The last interpretation of modular forms is perhaps best motivated by the classical view of modular forms as certain homogenous functions of lattices. Take for simplicity $\Gamma = \mathrm{SL}_2(\mathbb{Z})$.

To give an elliptic curve $E_{/\mathbb{C}}$ up to isomorphism is to give a lattice $\mathcal{L}_E$ in $\mathbb{C}$ up to homothety. Thus to give an elliptic curve with a non-vanishing differential $(E_{/\mathbb{C}}, \omega_0)$ is to give a unique lattice in $\mathbb{C}$, $\mathcal{L}_{E,\omega_0}$. Hence, a modular form f gives a function on lattices in $\mathbb{C}$. Since we have the relation

$$\mathcal{L}_{E,\lambda\omega_0} = \lambda\mathcal{L}_{E,\omega_0} \tag{4.18}$$

the property $f(E, \lambda\omega_0) = \lambda^{-k} f(E, \omega_0)$ is translated to the property $f(\lambda\mathcal{L}) = \lambda^{-k} f(\mathcal{L})$. I.e., f is a function of lattices that is homogenous of weight k. See also [**101**, Chapter VII], [**66**, Chapter 3, Section 2].

Let f be a modular form with respect to the group $\Gamma_0(N)$ for some N. Since the matrix $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ belongs to $\Gamma_0(N)$ it follows that $f(\tau + 1) = f(\tau)$. Therefore, f

has a Fourier expansion, called a *q-expansion* with respect to the variable $q = e^{2\pi i\tau}$

$$f = \sum_{n \geq -N} a_n q^n. \tag{4.19}$$

A similar expansion can be obtained in every cusp with respect to a suitable parameter. The holomorphic modular forms have q-expansions with $a_n = 0$ at every cusp if $n < 0$, and the cusp forms have q-expansions with $a_n = 0$ at every cusp if $n \leq 0$. We note that if we let $\mathbb{E}^k$ denote the unique extension of $\mathbb{E}^k$ to a sheaf on $X_0(N)(\mathbb{C}) = \Gamma_0(N)\backslash\mathcal{H}^*$ that is invertible at the cusps divisor **cusps**, then the holomorphic modular forms of weight k are just the sections $H^0(X_0(N)(\mathbb{C}), \mathbb{E}^k)$, while the cusp forms of weight k are sections of $H^0(X_0(N)(\mathbb{C}), \mathbb{E}^k(-\mathbf{cusps}))$.

5. Some Examples of Modular Forms

5.1. Level 1. Let k be an even integer, $k \geq 4$. The *Eisenstein series of weight* k is the complex valued function on the upper half plane

$$G_k^{\mathbb{Q}}(\tau) = \sum{}' \frac{1}{(m\tau + n)^k}. \tag{5.1}$$

The sum extending over all integers m, n such that $(m, n) \neq (0, 0)$. It is well known that this is a modular form of level one and weight k. See [**101**, Chapter VII]. We normalize $G_k^{\mathbb{Q}}$ such that the q-expansion starts with one and denote it then by $E_k^{\mathbb{Q}}$ or simply by E_k. Then the expansion of $E_k^{\mathbb{Q}}$ is given by (loc. cit.)

form	q-expansion
$E_4^{\mathbb{Q}}$	$1 + 240\sum_{n=1}^{\infty} \sigma_3(n)q^n$
$E_6^{\mathbb{Q}}$	$1 - 504\sum_{n=1}^{\infty} \sigma_5(n)q^n$
$E_8^{\mathbb{Q}}$	$1 + 480\sum_{n=1}^{\infty} \sigma_7(n)q^n$
$E_{10}^{\mathbb{Q}}$	$1 - 264\sum_{n=1}^{\infty} \sigma_9(n)q^n$
$E_{12}^{\mathbb{Q}}$	$1 + \frac{65520}{691}\sum_{n=1}^{\infty} \sigma_{11}(n)q^n$
$E_{14}^{\mathbb{Q}}$	$1 - 24\sum_{n=1}^{\infty} \sigma_{13}(n)q^n$

(5.2)

where $\sigma_r(n) = \sum_{d|n} d^r$, and in general

$$E_k(q) = 1 + 2\zeta(1-k)^{-1} \sum_{n=1}^{\infty} \sigma_{k-1}(n)q^n. \tag{5.3}$$

It is well known that the graded ring of modular forms of level 1 over the complex numbers, $\oplus_{k=0}^{\infty} M(k, \mathrm{SL}_2(\mathbb{Z}))$, is isomorphic to the free polynomial ring in two variables $\mathbb{C}[E_4, E_6]$ of weights 4 and 6 respectively. The modular form $\Delta = \frac{1}{1728}(E_4^3 - E_6^2)$ is a cusp form of weight 12 that induces an injection

$$M(k, \mathrm{SL}_2(\mathbb{Z})) \longrightarrow M(k+12, \mathrm{SL}_2(\mathbb{Z})), \tag{5.4}$$

whose image is of codimension one, consisting of all the cusp forms of weight $k+12$. Thus, for $k \geq 0$,

$$\dim\ M(k+12, \mathrm{SL}_2(\mathbb{Z})) = \dim\ M(k, \mathrm{SL}_2(\mathbb{Z})) + 1. \tag{5.5}$$

The dimensions for small weight are given in the following table:

(5.6)

weight	dimension	basis	weight	dimension	basis
0	1	1	12	2	$\Delta, E_{12}^{\mathbb{Q}}$
2	0	–	14	1	$E_{14}^{\mathbb{Q}}$
4	1	$E_4^{\mathbb{Q}}$	16	2	$\Delta E_4^{\mathbb{Q}}, E_{16}^{\mathbb{Q}}$
6	1	$E_6^{\mathbb{Q}}$	18	2	$\Delta E_6^{\mathbb{Q}}, E_{18}^{\mathbb{Q}}$
8	1	$E_8^{\mathbb{Q}}$	20	2	$\Delta E_8^{\mathbb{Q}}, E_{20}^{\mathbb{Q}}$
10	1	$E_{10}^{\mathbb{Q}}$	22	2	$\Delta E_{10}^{\mathbb{Q}}, E_{22}^{\mathbb{Q}}$

The modular form Δ has some important properties making it one of the key players in the theory. First, it has a product expansion

$$\Delta = q\prod_{n=1}^{\infty}(1-q^n)^{24}, \tag{5.7}$$

that shows that Δ does not vanish on the upper half plane, and has integral Fourier coefficients $\sum_{n=1}^{\infty}\tau(n)q^n$. The coefficients $\tau(n)$ are the famous Ramanujan τ-function. Second, it is the unique cusp form of its weight and thus an eigenvalue for all the Hecke operators. Third, it has the interpretation of being (up to a constant) the discriminant of the elliptic curve $\mathcal{E}_\tau$ associated to $\tau\in\mathcal{H}$ by means of the Weierstrass $\wp$-function (and that shows again that Δ does not vanish).

5.2. Higher level. Let f be a modular form of level 1 and weight k considered as a function on $\mathcal{H}$. Consider the function

$$\tau\mapsto g(\tau) := f(N\tau). \tag{5.8}$$

Let $\gamma=\begin{pmatrix} a & b\\ c & d\end{pmatrix}$ be a matrix in $\Gamma_0(N)$. Then $g(\gamma\tau)=f(N\gamma\tau)=f(N\gamma N^{-1}N\tau)$ $=f(\begin{pmatrix} a & Nb\\ c/N & d\end{pmatrix}N\tau)=(c\tau+d)^kf(N\tau)=(c\tau+d)^kg(\tau)$. Thus $g(\tau)$ is a modular form of level $\Gamma_0(N)$ and weight k.

Such modular forms are very important. Consider for example the modular function of level N given by

$$\tau\mapsto u(\tau) := \frac{\Delta(N\tau)}{\Delta(\tau)}. \tag{5.9}$$

The divisor of this function is supported at the cusps. In fact, for N prime there are two cusps c_1, c_2 on $X_0(N)$ and this function proves the Manin-Drinfeld theorem: *every divisor D of degree zero supported at the cusps is torsion. That is, there exists a positive integer n and a function h whose divisor is nD.*

EXERCISE 5.1. Bound the order of the divisor c_1-c_2. In fact, one can prove that the order is exactly the numerator of $(N-1)/12$ but this is harder. See [**89**, Section 3] and [**90**, Section 4] for generalizations.

In general, one considers the value of this function at a point τ. Since it has integral Fourier coefficients it extends to a function over the arithmetic scheme over $\mathbb{Z}[1/N]$ – the proper regular model of $X_0(N)$ – and still its divisor is supported at the cusps because the g.c.d. of the Fourier coefficients is 1. Thus, from the moduli interpretation we see that if the elliptic curve corresponding to τ is defined over $\mathbb{Q}^{alg}$ and has everywhere potential good reduction then $u(\tau)$ is at least an S-unit,

where S is the set of primes dividing N. Moreover, this unit lies in the moduli field of $\mathcal{E}_\tau$, which is equal to the residue field of τ as a point on the algebraic curve $X_0(N)$. In case τ corresponds to an elliptic curve with complex multiplication one can further find this moduli field and the Galois action via the theory of complex multiplication and Shimura's reciprocity law. Those units are called Siegel's units. . See [**66**, Chapter 12] and [**26**, Chapter 2.2].

It is not our intention to give a systematic introduction to the theory of elliptic modular forms, i.e. to modular forms for some congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$, but rather to give some interesting examples that connect to the material to be discussed in this book. We thus restrict our attention to two constructions of modular forms: Eisenstein series and theta series. We follow Ogg [**88**], where the reader can find proofs and more details.

Eisenstein series. Let $k \geq 3$. [2] Let N be a fixed positive integer and c, d be two integers. Generalizing the construction of the Eisenstein series of level one, one defines

$$G_k(\tau; c, d, N) = \sum_{\substack{m \equiv c \pmod N \\ n \equiv d \pmod N}}{}' \frac{1}{(m\tau + n)^k}. \tag{5.10}$$

It is easy to verify that $G_k(\tau; c, d, N)$ is a modular form of weight k and level $\Gamma(N)$. One can modify the construction and define the restricted Eisenstein series $G_k^*(\tau; c, d, N)$ for $(c, d, N) = 1$ as

$$G_k^*(\tau; c, d, N) = \sum_{\substack{m \equiv c \pmod N \\ n \equiv d \pmod N \\ (m,\, n) = 1}}{}' \frac{1}{(m\tau + n)^k}. \tag{5.11}$$

Also $G_k^*(\tau; c, d, N)$ is a modular form of weight k and level $\Gamma(N)$. The two kinds of Eisenstein series are related by explicit formulae ([**88**, IV-34]).

The restricted Eisenstein series $G_k^*(\tau; c, d, N)$ have the wonderful property that their q-expansion starts with zero at all cusps except for the cusp $-c/d$ where the q-expansion starts with a constant α that is independent of c and d. The restricted Eisenstein series are thus seen to generate the space of Eisenstein series.

Recall the notation $f|\gamma$:

$$(f|\gamma)(\tau) = (ad - bc)^{k/2}(c\tau + d)^{-k} f(\gamma\tau), \quad \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(\mathbb{R})^+. \tag{5.12}$$

We have with this notation

$$G_k^*(\tau; (c, d), N)|\gamma = G_k^*(\tau; (c, d)\gamma, N). \tag{5.13}$$

([**88**, IV-33]). Thus, from orbits of $\Gamma_0(N)$ in the set of pairs of residues (c, d) modulo N such that $(c, d, N) = 1$, we get Eisenstein series for $\Gamma_0(N)$.

[2] For level subgroups, unlike the case of $\mathrm{SL}_2(\mathbb{Z})$ they may be non-trivial modular forms of odd weight, though there are non of negative weight.

Theta series. Let r be an even integer and let

$$Q(x) = \frac{1}{2}\,{}^t x A x = \frac{1}{2}\sum_{i,j=1}^{r} a_{ij}x_i x_j \tag{5.14}$$

be a positive definite integral quadratic form. That is $A = (a_{ij})$ is a symmetric $r \times r$ matrix of integers whose diagonal elements are even, and A is positive definite. One defines the *determinant* of Q, denoted D, as the determinant of A, and Δ, the *discriminant* of Q, as $(-1)^{r/2}D$. The least positive integer N with $A^* = NA^{-1}$ integral symmetric matrix is called the *level* of Q, and the integral positive definite form it defines

$$Q^* = \frac{1}{2}\,{}^t x A^* x \tag{5.15}$$

is the *adjoint form* to Q. One always have $N|D|N^r$. See [**88**, Chapter VI].

The theta series associated to Q is defined as

$$\begin{aligned}\theta(\tau, Q) &= \sum_{v\in\mathbb{Z}^r} e^{2\pi i\cdot Q(v)\tau} \\ &= 1 + \sum_{n=1}^{\infty} a_Q(n) q^{n/2}, \quad q = e^{2\pi i\tau} \text{ (sic!)}.\end{aligned} \tag{5.16}$$

Here the coefficients $a_Q(n)$ are the representation numbers of Q, i.e., $a_Q(n)$ is the number of vectors $v \in \mathbb{Z}$ solving the equation $Q(x) = n$.

Let ϵ be the character $\epsilon(n) = \left(\frac{\Delta}{n}\right)$ (Jacobi symbol. See [**51**, Page 56]). Then the main theorem is that $\theta(\tau, Q)$ *is a modular form for* $\Gamma_0(n)$ *of character* ϵ *and weight* $r/2$. (This construction can be generalized using spherical functions. See [**88**, VI-10]).

It turns out that to get modular forms of level 1 the number of variables must be divisible by 8. See [**101**, Chapter VII, 6.5].

EXAMPLE 5.2.

1. Let $Q_N(x) = Nx^2$ corresponding to the matrix $A_N = 2N$. Note that $\theta(\tau, Q_N) = \theta(N\tau, Q_1)$ and

$$\theta(\tau, Q_N) = 1 + 2\sum_{n=1}^{\infty} q^{Nn^2/2}. \tag{5.17}$$

   We note that

$$\theta(\tau, Q_1) = 1 + 2\sum_{n=1}^{\infty} q^{n^2/2} = \Theta\begin{bmatrix}0\\0\end{bmatrix}(0, \tau), \tag{5.18}$$

   (Riemann's theta function, [**33**, Chapter VI]). The level and discriminant of Q_1 is 2 and we get that $\theta(\tau, Q_1)$ is a modular form of weight $1/2$, level $\Gamma_0(2)$ and trivial character.
2. Let $Q(x) = x_1^2 + x_2^2 + x_3^2 + x_4^2$ corresponding to the diagonal matrix $A = \mathrm{diag}[2, 2, 2, 2]$. It is a quadratic form in four variables, of level 2 and discriminant 16. It therefore defines a modular form of level $\Gamma_0(2)$, weight 2 and trivial character.

$$\theta(\tau, Q) = 1 + \sum_{n=1}^{\infty} a_n q^{n/2}, \tag{5.19}$$

where a_n is the number of expressions of n as a sum of four squares (negative numbers allowed). Note that

$$\theta(\tau, Q) = \theta(\tau, Q_1)^4. \tag{5.20}$$

A classical formula gives an elementary expression for the coefficients a_n:

PROPOSITION 5.3. (Jacobi) *The number of representation a_n of n as a sum of four squares is*

$$a_n = \begin{cases} 8\sum_{d|n} d & n \text{ odd} \\ 24\sum_{d|n, d\equiv 1 \ (2)} d & n \text{ even} \end{cases}. \tag{5.21}$$

Two strategies for proving this Proposition are as follows: (i) Observe that since $X_0(2)$ is of genus zero, there is a unique modular form of level 2 up to a scalar. Thus, one constructs a modified Eisenstein series of level two and compares coefficients. See [**80**, Chapter 1, Section 15]. (ii) Use the product expansion of $\theta(\tau, Q_1)$ to derive the formulae directly (loc. cit.).

Connection to supersingular elliptic curves. Let p be a prime number and consider the set of isomorphism classes of supersingular elliptic curves over $\overline{\mathbb{F}_p}$. There are $\hbar = g + 1$ of them where g is equal the genus of $X_0(p)$:

$$g = \begin{cases} \frac{p+1}{12} - \frac{1+\left(\frac{-1}{p}\right)}{4} - \frac{1+\left(\frac{-3}{p}\right)}{3} & p \neq 2 \\ 0 & p = 2 \end{cases}. \tag{5.22}$$

EXERCISE 5.4. Prove formula (5.22) from Hurwitz genus formula using that $\mathrm{SL}_2(\mathbb{Z})\backslash\mathcal{H}$ has genus zero.

Let E be an elliptic curve over $\overline{\mathbb{F}_p}$. We recall here various equivalent ways of saying that E is supersingular: (i) E has no physical p-torsion: $E[p](\overline{\mathbb{F}_p}) = \{0\}$; (ii) there exists an embedding of the group scheme α_p into E; (iii) there does not exist an embedding of μ_p into E; (iv) the Frobenius morphism $\mathrm{Fr}_E : E \longrightarrow E^{(p)}$ is purely inseparable; (v) the multiplication by p map $[p] : E \longrightarrow E$ is purely inseparable; (vi) the ring of endomorphisms of E is isomorphic to a (maximal) order of the quaternion algebra $B_{p,\infty}$ – the quaternion algebra over $\mathbb{Q}$ ramified at the two places p and ∞; (vii) for n large enough Fr_E^n is an endomorphism of E and its characteristic polynomial has Newton polygon w.r.t. the p-adic valuation which is a straight line; (viii) the formal group of E has height 2.

See [**107**, V.3] for some of the equivalences. The others you can do once you read Chapter 6 Section 7 and Appendix A. We recall that furthermore the j-invariant of a supersingular curve is in $\mathbb{F}_{p^2}$;

There is an interesting characteristic p method to prove formula (5.22). One considers the $\Gamma_0(p)$-level moduli problem of parameterizing elliptic curves with a given subgroup of order p. One proves that there exists a coarse moduli scheme for this moduli problem, which is a regular scheme that is a *flat* relative curve $Y_0(p)$ over $\mathbb{Z}$. Over the complex numbers $Y_0(p)(\mathbb{C}) \cong \Gamma_0(p)\backslash\mathcal{H}$.

To actually have a proper morphism $X_0(p) \longrightarrow \mathrm{Spec}(\mathbb{Z})$, one needs to throw in the cusps. The cusps may also be given a moduli interpretation using the concept of generalized elliptic curves. See [**22**]. This interpretation shows that

there are two cusps lying over the unique cusp of $X_0(1)$ (the j-line). One is unramified, and the other has ramification of order p, and that moreover this persists in characteristic p. One proves that $X_0(p)$ is a regular integral scheme and the morphism $X_0(p) \longrightarrow \mathrm{Spec}(\mathbb{Z})$ is *proper* and *flat*. Over the complex numbers $X_0(p)(\mathbb{C}) \cong \Gamma_0(p)\backslash\mathcal{H}^*$.

One goes on to prove that in characteristic p, the reduction of $X_0(p)$ consists of two components that intersect transversely precisely above the supersingular points of $X_0(p)$. See Figure ***

Figure 2.

The proof is based on the construction of two "sections"

$$X_0(1) \pmod p \rightrightarrows X_0(p) \pmod p. \tag{5.23}$$

One section is a true section and is obtained by sending an elliptic curve E to the pair $(E, \mathrm{Ker}(\mathrm{Fr}_E))$ (note that if E is ordinary $\mathrm{Ker}(\mathrm{Fr}_E)$ is a form of μ_p) and the other is obtained by sending E to $(E, \mathrm{Ker}(\mathrm{Ver}_E))$, where $\mathrm{Ver}_E : E \longrightarrow E^{(1/p)}$ (note that if E is ordinary $\mathrm{Ker}(\mathrm{Ver}_E)$ is a form of $\mathbb{Z}/p\mathbb{Z}$). Leaving aside the problem of non-integral base schemes, there is still the problem that Ver_E is usually defined only after a base change. See Appendix A for groups schemes and the morphisms Fr, Ver.

Thus the true state of affairs is that $X_0(p) \pmod p$ is composed of two components that we denote $X_0(p)^{\mu_p}$ and $X_0(p)^{\mathbb{Z}/p\mathbb{Z}}$. Those two components intersect precisely over the supersingular points. The map

$$X_0(p)^{\mu_p} \longrightarrow X_0(1) \pmod p, \tag{5.24}$$

is an isomorphism, while the map

$$X_0(p)^{\mathbb{Z}/p\mathbb{Z}} \longrightarrow X_0(1) \pmod p, \tag{5.25}$$

is purely inseparable of degree p. Thus, both components have genus zero, and the total genus of $X_0(p) \pmod p$ (which, courtesy of flatness, is also the genus of *every* fiber of $X_0(p) \longrightarrow \mathrm{Spec}(\mathbb{Z})$) is the genus of the intersection graph of the components. This graph consists of two vertices, with $\hbar$ edges between them, and has thus genus $\hbar - 1$.

Fix representatives for the supersingular elliptic curve $E_1, \ldots, E_\hbar$. For every i and j we can consider the abelian group

$$I_{i,j} = \mathrm{Hom}_{\overline{\mathbb{F}_p}}(E_i, E_j). \tag{5.26}$$

It is a left $\mathrm{End}(E_j)$ module and a right $\mathrm{End}(E_i)$ module. The function $f \mapsto \deg(f)$ is a positive definite integral quadratic form $Q_{i,j}$ on $I_{i,j}$ in four variables. One

defines the theta series

$$\theta(\tau, Q_{i,j}) := \frac{1}{|\mathrm{Aut}(E_j)|} \sum_{b \in I_{i,j}} q^{\deg(b)}. \tag{5.27}$$

The main theorem is that *the $\theta(\tau, Q_{i,j})$ are modular forms of weight 2 and level $\Gamma_0(p)$ (with trivial character) that span the space of modular forms of weight 2 and level $\Gamma_0(p)$.*

We refer the interested reader to [**42**], [**70**], [**96**] for more on this fascinating story. The proof that the level and character are as stated follows of course from the theorem of Theta series we explained above and is in [**96**, Theorem 2.14].

Example 5.5. There is a unique supersingular elliptic curve E over a field of characteristic two. Its endomorphism ring is the ring $\mathbb{Z}[i, j, (1+i+j+k)/2]$ in the rational Hamilton quaternions $\mathbb{H}$, where following classical notation we write the Hamilton quaternions as $\mathbb{Q} \oplus \mathbb{Q}i \oplus \mathbb{Q}j \oplus \mathbb{Q}k$ with the relations

$$-1 = i^2 = j^2 = k^2, \;\; ij = k = -ji. \tag{5.28}$$

This ring is the free $\mathbb{Z}$ module on $1, i, j$ and $(1+i+j+k)/2$. The norm of an element $a + bi + cj + dk$ is $a^2 + b^2 + c^2 + d^2$, and is equal to its degree as an endomorphism of E. We see that up to a scalar, the theta function derived from E is the theta function of the form

$$Q(a + bi + cj + dk) = a^2 + b^2 + c^2 + d^2. \tag{5.29}$$

We arrive at Example 5.2, 2.

Exercise 5.6. The situation above can be made explicit. We leave the verifications of details to the reader.

The equation for E may be chosen as

$$y^2 + y = x^3. \tag{5.30}$$

The substitutions

$$x \mapsto u^2 x + s^2, \quad y \mapsto y + su^2 x + t, \tag{5.31}$$

give an automorphism of E if and only if the following equations are satisfied:

$$u^3 = 1, \quad s^4 + s = 0, \quad t^2 + t + s^6 = 0. \tag{5.32}$$

We write this automorphism as $(u, s, t)x, (u, s, t)y$. The group of automorphisms is visibly of order 24. Now,

$$(u_1, s_1, t_1)(u, s, t) = (u_1 u, s_1 + u_1 s, t_1 + t + u_1^2 s_1 s^2). \tag{5.33}$$

This shows that there is a unique element of order two. It is $(1, 0, 1)$ and we denote it by -1. The notation is justified since in its action on E it is equal to the automorphism $[-1]$, acting by $(x, y) \mapsto (x, y + 1)$. Handy formulas are

$$(u, s, t)^2 = (u^2, s(1 + u), us^3), \quad -(u, s, t) = (u, s, t + 1). \tag{5.34}$$

We deduce that there are precisely six elements of order four. They are the elements of the set $\{(1, s, t) : s^3 = 1, \; t^2 + t + 1 = 0\}$.

Let s_1, s_2 and s_3 be the solutions to $s^3 = 1$ and let t_1, t_2 be the solutions to $t^2 + t + 1 = 0$. We may assume that $s_1 s_2^2 = t_1$. With these conventions we let

$$i := (1, s_1, t_1), \quad j := (1, s_2, t_1), \quad k := (1, s_3, t_1). \tag{5.35}$$

The usual relations are satisfied:

$$i^2 = j^2 = k^2 = -1, \quad ij = -ji = k. \tag{5.36}$$

It follows that $\{\pm 1, \pm i, \pm j, \pm k\}$ is Q_8 – the group of quaternions of order 8. The cyclic group C_3 of order 3 given by $\{(u, 0, 0) : u^3 = 1\}$ acts by conjugation on Q_8, permuting cyclically the elements i, j, k, and

$$\mathrm{Aut}(E) \cong C_3 \ltimes Q_8. \tag{5.37}$$

6. Abelian Varieties over $\mathbb{C}$

In this section we describe the theory of abelian varieties over the complex numbers. We often refer to the excellent book by Lange and Birkenhake [**67**]. The reader may also consult [**109**], [**16**] and [**75**] and [**85**]. We found it necessary, regretfully, to omit most proofs in order to keep these lecture notes to a reasonable size. Nevertheless, we go into detail in describing abelian varieties over $\mathbb{C}$ in hope of arming the novice with some intuition when facing the much more strenuous presentation of abelian varieties (or schemes) over arbitrary fields in later sections.

We assume some familiarity with Lie groups. Specifically, with the definition of a complex Lie group as a analytic manifold with a group structure such that the group operations are analytic, and with the exponential map that we recall below.

Definition 6.1. Let k be a field. An *abelian variety* $\mathcal{A}$ over k is a projective, connected algebraic group.

We consider the case $k = \mathbb{C}$. Let V denote a g-dimensional vector space over $\mathbb{C}$ and Λ a lattice in V (i.e. a subgroup of rank $2g$ such that $\Lambda \otimes \mathbb{R} \cong V$ under the map $\lambda \otimes \alpha \mapsto \alpha\lambda$). A *complex torus* is a complex Lie group isomorphic to $X = V/\Lambda$ for some lattice Λ. Note that X is compact.

Remark 6.2. Do not confuse a complex torus with an algebraic torus as in Section 1.

Theorem 6.3. *Any connected compact complex Lie group X is a complex torus.*

Proof. We first prove that
•X is abelian.
Consider the commutator map

$$X \times X \longrightarrow X, \tag{6.1}$$

$$(x, y) \mapsto [x, y] = xyx^{-1}y^{-1}. \tag{6.2}$$

Let U be an open neighborhood of 1. Let $x \in X$. Since $[x, 1] = 1 \in U$ and $[\cdot, \cdot]$ is continuous, for every $x \in X$, we have open sets V_x and W_x, with $x \in V_x$ and $1 \in W_x$, such that $[V_x, W_x] \subset U$. Since X is compact,

$$X = \cup_{x \in X} V_x = \cup_{i=1,\cdots,n} V_{x_i} \quad (\text{for some } x_1, \cdots, x_n). \tag{6.3}$$

Let $W = \cap_{i=1}^n W_{x_i}$, an open neighborhood of 1. Then

$$[X, W] \subset U. \tag{6.4}$$

We may assume that there exists an isomorphism $U \cong \mathbb{C}^g$. Hence, for every $w \in W$ we get an analytic function

$$X \longrightarrow \mathbb{C}^g \quad ; \quad x \mapsto [x, w]. \tag{6.5}$$

This function must be constant since X is compact. Hence, for every $w \in W$, we have $[X, w] = 1$ and thus

$$[X, W] = 1. \tag{6.6}$$

The set W is open and contains the identity element, so it generates an open subgroup Q of X that has the property $[Q, X] = 1$; but in a topological group, any open subgroup is also closed. Thus, $Q = X$ is abelian.

• It follows from the theory of Lie groups that for every complex Lie group X, there exists a unique holomorphic map

$$\mathbb{C} \times V \longrightarrow X \quad ; \quad (t, v) \mapsto \phi_v(t), \tag{6.7}$$

where $V = \mathfrak{t}_{X,1}$ is the tangent space at 1, such that

1. $\phi_v(t)$ is a holomorphic homomorphism in t: $\phi_v(t_1 + t_2) = \phi_v(t_1)\phi_v(t_2)$. In our case, since X is abelian, we write $\phi_v(t_1 + t_2) = \phi_v(t_1) + \phi_v(t_2)$.
2. $T\phi_v(t)$ takes the unit tangent vector to $\mathbb{C}$ at zero to v (i.e. $T\phi_v(\frac{d}{dt}|_0) = v$).

One defines the *exponential map*

$$e : V \longrightarrow X \tag{6.8}$$

by

$$e(v) = \phi_v(1). \tag{6.9}$$

Figure 3.

The two maps $\mathbb{C} \longrightarrow X$ given by

$$t \mapsto \phi_{sv}(t) \quad ; \quad t \mapsto \phi_v(st) \quad (s \text{ fixed }) \tag{6.10}$$

are holomorphic homomorphisms with derivatives at zero equal to sv. Thus, by uniqueness,

$$\phi_{sv}(t) = \phi_v(st), \tag{6.11}$$

and therefore

$$e(sv) = \phi_v(s). \tag{6.12}$$

Consider the map

$$f : \mathbb{C} \longrightarrow X \tag{6.13}$$

given by

$$t \mapsto e(tx) + e(ty) = \phi_x(t) + \phi_y(t) \tag{6.14}$$

for some fixed $x, y \in X$. Then

$$\begin{aligned} f(t_1+t_2) &= e((t_1+t_2)x) + e((t_1+t_2)y) \\ &= \phi_x(t_1+t_2) + \phi_y(t_1+t_2) \\ &= \phi_x(t_1) + \phi_x(t_2) + \phi_y(t_1) + \phi_y(t_2) \\ &= (\phi_x(t_1) + \phi_y(t_1)) + (\phi_x(t_2) + \phi_y(t_2)) \\ &= f(t_1) + f(t_2) \end{aligned} \tag{6.15}$$

Therefore f is a holomorphic homomorphism with derivative at 0 equal to $x+y$ (use the chain rule). Thus, $f(t) = \phi_{x+y}(t)$. Plugging in $t = 1$, we get

$$e(x) + e(y) = e(x+y); \tag{6.16}$$

so $e : V \longrightarrow X$ is a homomorphism. The image of e is an open subgroup (it contains a neighborhood of 1) and since X is connected , e is surjective. Having X compact gives us that the kernel is a lattice Λ, since lattices are the only discrete subgroups of vector spaces with compact quotient. □

COROLLARY 6.4. *Every abelian variety over $\mathbb{C}$ is a complex torus, hence the group law is commutative.*

COROLLARY 6.5. *Let X be a complex torus. Then*

$$\begin{aligned} X[n] :&= \{x | x \in X : nx = 0\} \\ &\cong \frac{1}{n}\Lambda/\Lambda \\ &\cong (\mathbb{Z}/n\mathbb{Z})^{2g}. \end{aligned} \tag{6.17}$$

COROLLARY 6.6. *Once we fix isomorphisms $\Lambda_i \cong \mathbb{Z}^{2g_i}$, we have an isomorphism of groups*

$$\mathrm{Hom}(X_1, X_2) \cong \{M \in M_{g_1 \times g_2}(\mathbb{C}) : M\Lambda_1 \subset \Lambda_2\} \hookrightarrow \{M \in M_{2g_1 \times 2g_2}(\mathbb{Z})\}. \tag{6.18}$$

In particular, $\mathrm{Hom}(X_1, X_2)$ *is a torsion free abelian group of rank* $\leq 4g_1g_2$.

COROLLARY 6.7. *Complex compact connected g-dimensional Lie groups are parameterized by:*

$$\mathrm{GL}_g(\mathbb{C}) \backslash \mathrm{GL}_{2g}(\mathbb{R}) / \mathrm{GL}_{2g}(\mathbb{Z}). \tag{6.19}$$

PROOF. Any two lattices in $\mathbb{R}^{2g} \cong \mathbb{C}^g$ are equivalent under $\mathrm{GL}_{2g}(\mathbb{R})$. The equivalence is obtained by first choosing a basis for the lattice (this gives the matrix in $\mathrm{GL}_{2g}(\mathbb{R})$. Changing the choice of basis amounts to moding out by $\mathrm{GL}_{2g}(\mathbb{Z})$. Thus, $\mathrm{GL}_{2g}(\mathbb{R})/\mathrm{GL}_{2g}(\mathbb{Z})$ is a bijection with lattices in $\mathbb{R}^{2g}$. Moding out by $\mathrm{GL}_g(\mathbb{C})$ (embedded in $\mathrm{GL}_{2g}(\mathbb{R})$ by the isomorphism $\mathbb{R}^{2g} \cong \mathbb{C}^g$) amounts to the fact that two complex tori $\mathbb{C}^g/\mathcal{L}_1$ and $\mathbb{C}^g/\mathcal{L}_2$ are isomorphic iff there exists a matrix $M \in \mathrm{GL}_g(\mathbb{C})$ such that $M\mathcal{L}_1 = \mathcal{L}_2$. □

REMARK 6.8. While every g-dimensional abelian variety over $\mathbb{C}$ is a complex torus the converse does not hold if $g > 1$. In fact, the moduli of complex tori is g^2 dimensional, while that of abelian varieties is $g(g+1)/2$ dimensional. We remarked before (see Section 2) that the moduli of curves of genus g is $3g-3$ for $g > 1$. We sketch here the argument for abelian varieties following the same reasoning of Kodaira-Spencer we used for curves.

The Kodaira-Spencer theory dictates that the universal $k[\epsilon]$-deformation, i.e., first order deformation, of an abelian variety X is given by $H^1(X, \mathfrak{t}_X)$, where $\mathfrak{t}_X$ is the tangent sheaf to X. I.e., the sheaf of derivations. Moreover, the cup product map "evaluating the differential on a derivation"

$$H^1(X, \mathfrak{t}_X) \times H^0(X, \Omega^1_X) \longrightarrow H^1(X, \mathcal{O}_X), \tag{6.20}$$

yields an isomorphism

$$H^1(X, \mathfrak{t}_X) = \mathrm{Hom}(H^0(X, \Omega^1_X), H^1(X, \mathcal{O}_X)). \tag{6.21}$$

Now, for abelian variety one can identify $H^0(X, \Omega^1_X)$ with the cotangent space at zero $\mathfrak{t}^*_{X,0}$, and $H^1(X, \mathcal{O}_X)$ with the tangent space of the dual abelian variety $\mathfrak{t}_{X^t,0}$ discussed extensively below (Section 6.2). Thus,

$$H^1(X, \mathfrak{t}_X) = \mathfrak{t}_{X,0} \otimes \mathfrak{t}_{X^t,0}. \tag{6.22}$$

That shows that the first order deformations are of dimension g^2.

Moreover, if X is principally polarized (Section 6.2) [3] then this polarization λ gives a map λ_* from $\mathfrak{t}_{X,0}$ to $\mathfrak{t}_{X^t,0}$. The theory then identifies the deformations preserving the polarization with the symmetric tensors of $\mathfrak{t}_{X,0} \otimes \mathfrak{t}_{X^t,0}$ (under the involution induced by $v \otimes w \mapsto \lambda_*^{-1}(w) \otimes \lambda_*(v)$) giving moduli of dimension $g(g+1)/2$.

EXERCISE 6.9. Deduce that if a complex compact curve is a complex Lie group then it must be of genus 1. Prove, using just topological tools, that a closed real two dimensional manifold is a topological group if and only if it has genus 1.

EXERCISE 6.10. Prove that any holomorphic differential on an abelian variety X/k is translation invariant (and hence non-vanishing). Conclude the identification $H^0(X, \Omega^1_{X/k}) = \mathfrak{t}^*_{X,0}$ used above.

6.1. The Appell-Humbert theorem. Let's recall Chow's theorem.

THEOREM 6.11. *Let P be a complete algebraic variety over $\mathbb{C}$ and $\mathcal{Z} \subset P$ a closed analytic subset of P; then there is an algebraic sub-variety Z of P such that $Z(\mathbb{C}) \cong \mathcal{Z}$, as complex analytic spaces.*

PROOF. See [**41**, Page 167]. □

DEFINITION 6.12. Let X be a complete analytic variety. A line bundle $\mathcal{L} \longrightarrow X$ is *ample* if $\exists k > 0$ such that $\mathcal{L}^{\otimes k}$ is *very ample*, i.e., if we have an embedding:

$$X \hookrightarrow \mathbb{P}(H^0(X, \mathcal{L}^k)), \tag{6.23}$$

$$x \mapsto \{\text{hyperplane of sections vanishing at } x\}. \tag{6.24}$$

If one chooses a basis $s_0, \ldots, s_n$ to $\Gamma(X, \mathcal{L}^{\otimes k})$, we can also write such an embedding as

$$x \mapsto (s_0(x) : \cdots : s_n(x)). \tag{6.25}$$

COROLLARY 6.13. *A complex torus is an abelian variety if and only if it has an ample line bundle $\mathcal{L}$.*

[3] More generally, in any characteristic in fact, if X has a polarization that is a separable map.

We would like to understand when is the complex torus $\mathbb{C}^g/\Lambda$ an abelian variety. Given the corollary, the main problem we are faced with is:

(6.26) *"When does $\mathbb{C}^g/\Lambda$ have an ample line bundle?, and even more, what are all the line bundles on $\mathbb{C}^g/\Lambda$?"*

The Appell-Humbert supplies a complete answer to those questions.

First, let us make a digression on *factors of automorphy* (compare Definition 4.1): Let $X = \mathbb{C}^g/\Lambda$ be a complex torus, $\pi : \mathbb{C}^g \longrightarrow X$ the natural projection and $\mathcal{L}$ a line bundle on X. Because $\mathbb{C}^g$ is simply connected, every line bundle on it is trivial, i.e. $\pi^*\mathcal{L} \cong \mathbb{C}^g \times \mathbb{C}$. Let T_λ be the translation map on $\mathbb{C}_g : T_\lambda(x) = x + \lambda$. For every $\lambda \in \Lambda$, we have

$$T_\lambda^*\pi^*\mathcal{L} = (\pi \circ T_\lambda)^*\mathcal{L} = \pi^*\mathcal{L}. \tag{6.27}$$

But

$$(T_\lambda^*(\pi^*\mathcal{L}))_x = \pi^*\mathcal{L}_{x+\lambda}, \tag{6.28}$$

by definition of the pullback of a line bundle. This implies the existence of a holomorphic non-vanishing map

$$j(\lambda) : \mathbb{C}^g \longrightarrow \mathbb{C}^\times \tag{6.29}$$

such that

$$j(\lambda_1)(\lambda_2 + v) \cdot j(\lambda_2)(v) = j(\lambda_1 + \lambda_2)(v) \tag{6.30}$$

(it is the same compatibility requirement as in Section 4). We let Λ act on holomorphic functions f by ${}^\lambda f(v) = f(\lambda + v)$. We see that

$$j(\lambda_1 + \lambda_2) = {}^{\lambda_2}j(\lambda_1) \cdot j(\lambda_2) \tag{6.31}$$

It follows that

$$j \in Z^1(\Lambda, \underline{\mathcal{O}}^\times(\mathbb{C}^g)), \tag{6.32}$$

where $\underline{\mathcal{O}}^\times(\mathbb{C}^g)$ denote the group of non-vanishing holomorphic functions on $\mathbb{C}^g$. Conversely, any $j \in Z^1(\Lambda, \underline{\mathcal{O}}^\times(\mathbb{C}^g))$ induces a line bundle on $\mathbb{C}^g/\Lambda$ whose isomorphism class depends only on the image of j in $H^1(\Lambda, \underline{\mathcal{O}}^\times(\mathbb{C}^g))$

One can show the following

PROPOSITION 6.14. *There exists canonical isomorphisms:*

(6.33)

$$\begin{aligned} \mathrm{Pic}(X) :&= \{ \textit{ isomorphism classes of line bundles on } X\} \\ &\cong H^1(X, \mathcal{O}_X^\times) && \textit{(Sheaf cohomology)} \\ &\cong H^1(\Lambda, \underline{\mathcal{O}}^\times(\mathbb{C}^g)) && \textit{(Group cohomology).} \end{aligned}$$

PROOF. See [**67**, Page 24]. □

Thus, we are only left with the task of describing this latter group !

DEFINITION 6.15. A *Riemann form* on $\mathbb{C}^g$ *with respect to* Λ is a hermitian form $H : \mathbb{C}^g \times \mathbb{C}^g \longrightarrow \mathbb{C}$ such that $E = \mathrm{Im}\ H$ (meaning $E(u, v) = \mathrm{Im}(H(u, v))$) satisfies

$$E : \Lambda \times \Lambda \longrightarrow \mathbb{Z}. \tag{6.34}$$

Recall that a hermitian form is $\mathbb{C}$-linear in the first variable, $\overline{\mathbb{C}}$-linear in the second variable, and $H(v,w) = \overline{H(w,v)}$.

DEFINITION 6.16. A *semi-character* with respect to a Riemann form H is a map

$$\chi : \Lambda \longrightarrow \mathbb{C}_1 := \{z \in \mathbb{C} : |z| = 1\} \tag{6.35}$$

such that

$$\chi(\lambda_1 + \lambda_2) = \chi(\lambda_1)\chi(\lambda_2)\exp(\pi i E(\lambda_1, \lambda_2)). \tag{6.36}$$

Note that although χ is not necessarily a character, χ^2 *is* a character.

EXERCISE 6.17. 1. Let $\mathcal{L}$ be a lattice in $\mathbb{C}$. Find all Riemann forms and quasi-characters on $\mathcal{L}$. In particular, deduce that a non-trivial Riemann form always exists.
2. Find a lattice in $\mathbb{C}^2$ that has no non-trivial Riemann form.

We define a group structure on the set

$$\mathcal{G} = \{(H,\chi) : H \text{ a Riemann form with respect to } \Lambda, \ \chi \text{ a semi-character with respect to } H\}. \tag{6.37}$$

by putting

$$(H_1,\chi_1) + (H_2,\chi_2) = (H_1 + H_2, \chi_1\chi_2). \tag{6.38}$$

An element $(H,\chi) \in \mathcal{G}$ induces a line bundle on $\mathbb{C}^g/\Lambda$ denoted $\mathcal{L}_{(H,\chi)}$, defined by the factor of automorphy $j_{(H,\chi)}$:

$$j_{(H,\chi)}(\lambda)(v) = \chi(\lambda)\exp(\pi H(v,\lambda) + \frac{\pi}{2}H(\lambda,\lambda)), \quad v \in \mathbb{C}^g, \lambda \in \Lambda. \tag{6.39}$$

THEOREM 6.18. (Appell-Humbert) *There exists an isomorphism of groups,*

$$\mathcal{G} \cong \mathrm{Pic}(X), \ \ (H,\chi) \mapsto \mathcal{L}_{(H,\chi)}. \tag{6.40}$$

In particular,

$$\mathcal{L}_{(H_1,\chi_1)} \otimes \mathcal{L}_{(H_2,\chi_2)} \cong \mathcal{L}_{(H_1+H_2,\chi_1\chi_2)}. \tag{6.41}$$

Moreover, $\mathcal{L}_{(H,\chi)}$ is ample if and only if H is positive definite .

PROOF. See [**67**], Page 32 and Proposition 5.2, Page 86. □

COROLLARY 6.19. *A complex torus $X = \mathbb{C}^g/\Lambda$ is an abelian variety if and only if there exists a positive definite Riemann form H with respect to Λ.*

REMARK 6.20. 1. One can show E is alternating (i.e. skew commutative) and satisfies the identity

$$E(ix, iy) = E(x,y). \tag{6.42}$$

Conversely, given an alternating form E satisfying Equations (6.42) and (6.34) it is easy to check that the function $H(x,y) = E(ix,y) + iE(x,y)$ is a Riemann form.

2. One may obtain semi-characters as follows: Assume E is non-degenerate. It is well-known (Elementary Divisors Theorem) that there exists a basis

$$x_1, \dots, x_g, y_1, \dots, y_g, \tag{6.43}$$

such that E is given by the matrix:

$$\begin{pmatrix} 0 & D \\ -D & 0 \end{pmatrix}, \tag{6.44}$$

where 0 stands for the $g \times g$ zero matrix and D is a $g \times g$ diagonal matrix with diagonal entries $d_i \in \mathbb{N}$, such that $d_1 | d_2 | \cdots | d_g$. Put

$$\Lambda_1 = \mathrm{Span}_{\mathbb{Z}}\{x_i : i = 1, \dots, g\}, \;\; \Lambda_2 = \mathrm{Span}_{\mathbb{Z}}\{y_i : i = 1, \dots, g\}. \tag{6.45}$$

Clearly

$$\Lambda = \Lambda_1 \oplus \Lambda_2. \tag{6.46}$$

Then $\mathbb{C}^g = \mathbb{R}\Lambda_1 \oplus \mathbb{R}\Lambda_2$, and we write the decomposition of a vector $v \in \mathbb{C}^g$ accordingly as

$$v = v_1 + v_2. \tag{6.47}$$

Put

$$\chi(v) = \exp[\pi i \cdot E(v_1, v_2)]. \tag{6.48}$$

Then

$$\begin{aligned} \chi(v+w) &= \exp[\pi i \cdot E(v_1 + w_1, v_2 + w_2)] \\ &= \exp[\pi i (E(v_1, v_2) + E(w_1, w_2) + E(v_1, w_2) + E(w_1, v_2))] \\ &= \chi(v) \cdot \chi(w) \cdot \exp[\pi i (E(v_1, w_2) + E(w_1, v_2))] \\ &= \chi(v) \cdot \chi(w) \cdot \exp[\pi i (E(v, w))], \end{aligned} \tag{6.49}$$

because Λ_1, Λ_2 are isotropic with respect to E.

Remark 6.21. Given $\mathcal{L} \in \mathrm{Pic}(X)$, we may ask what is the associated Riemann or hermitian form?

Consider the exact sequence of sheaves on X:

$$0 \longrightarrow \mathbb{Z} \longrightarrow \underline{\mathcal{O}} \xrightarrow{e^{2\pi i(-)}} \underline{\mathcal{O}}^{\times} \longrightarrow 0. \tag{6.50}$$

The associated cohomology sequence gives a map

$$\mathrm{Pic}(X) := H^1(X, \underline{\mathcal{O}}_X^{\times}) \longrightarrow H^2(X, \mathbb{Z}). \tag{6.51}$$

Under this map, the line bundle $\mathcal{L}$ is sent to an element $c_1(\mathcal{L}) \in H^2(X, \mathbb{Z})$ called the *first Chern class* of $\mathcal{L}$. However,

$$H^2(X, \mathbb{Z}) = \wedge^2 H^1(X, \mathbb{Z}). \tag{6.52}$$

That is, if $X = \mathbb{C}^g/\Lambda$ then $H^2(X, \mathbb{Z})$ consists of alternating $\mathbb{Z}$-valued bilinear forms on $H_1(X, \mathbb{Z}) = \Lambda$ (To an element $\sum \phi_i \wedge \psi_i$ associate the alternating form $(v, w) \mapsto \sum \phi(v) - \psi(w)$. Thus $c_1(\mathcal{L})$ is such a bilinear form E which is precisely $\mathrm{Im}H$ where H is the Riemann form associated to $\mathcal{L}$.

6.2. The dual abelian variety.

DEFINITION 6.22. Let X be a complex torus. The Néron-Severi group of X is defined as the image of the homomorphism $H^1(X, \mathcal{O}_X^\times) \longrightarrow H^2(X, \mathbb{Z})$. It is denoted by $NS(X)$. We put $NS^0(X) = NS(X) \otimes \mathbb{Q}$.

REMARK 6.23. The Néron-Severi group is also the group of divisors on X modulo algebraic equivalence.[4] Using our knowledge on complex tori we see that for $X = V/\Lambda$, $NS(X)$ can be identified with the group of Riemann forms H with respect to Λ. Since it is a subgroup of $H^2(X, \mathbb{Z}) = H^1(X, \mathbb{Z}) \wedge H^1(X, \mathbb{Z})$ (via $H \mapsto \mathrm{Im}(H)$), it is a free $\mathbb{Z}$-module of rank $\leq g(2g-1)$, where $g = \dim(X)$. See [**47**, p.447] for further details. We let $NS(X)^+$ denote the elements corresponding to positive definite Riemann forms with respect to Λ.

DEFINITION 6.24. The subgroup of $\mathrm{Pic}(X)$ given by

$$\{(0, \chi) : \chi : \Lambda \longrightarrow \mathbb{C}_1 \text{ is a homomorphism}\} \tag{6.53}$$

is called the dual abelian variety. It is also denoted by $\mathrm{Pic}^0(X)$, X^t, or $X^\vee$.

These groups are related by an exact sequence:

$$0 \longrightarrow \mathrm{Pic}^0(X) \longrightarrow \mathrm{Pic}(X) \longrightarrow NS(X) \longrightarrow 0 \tag{6.54}$$

The dual abelian variety Pic^0 carries a canonical complex structure, as we will see shortly. To be precise, that canonical complex structure is part of the definition of the dual abelian variety. We ease notation and write $\mathrm{Pic}^0(X) = \{\chi | \chi : \Lambda \longrightarrow \mathbb{C}_1\}$.

Consider the complex torus $X = V/\Lambda$, where V is g-dimensional vector space, and Λ is a full lattice. Put

$$\Omega = \Omega(X) = \mathrm{Hom}_{\overline{\mathbb{C}}}(V, \mathbb{C}), \tag{6.55}$$

where $\overline{\mathbb{C}}$ means $\mathbb{C}$-anti-linear; Ω is a complex vector space of dimension g and

$$f \in \Omega \iff f(v_1) + f(v_2) = f(v_1 + v_2), \quad f(\lambda v) = \overline{\lambda} f(v). \tag{6.56}$$

The pairing

$$\Omega \times V \longrightarrow \mathbb{C}, \quad (f, v) \mapsto f(v) \tag{6.57}$$

is bi-additive, $\mathbb{C}$-linear in the first variable, and $\overline{\mathbb{C}}$-linear in the second variable. We have an isomorphism

$$\Omega \xrightarrow{\cong} \mathrm{Hom}_{\mathbb{R}}(V, \mathbb{R}) \tag{6.58}$$

given by

$$l \longmapsto \mathrm{Im}\ l. \tag{6.59}$$

The inverse is given by $k \longmapsto l$, where $l(v) = -k(iv) + ik(v)$. Thus,

$$\Omega \times V \longrightarrow \mathbb{R}, \quad (l, v) \longmapsto \mathrm{Im}(l(v)) \tag{6.60}$$

is a perfect $\mathbb{R}$-linear pairing.

Put

$$\hat{\Lambda} = \{l \in \Omega : \mathrm{Im}\ l(v) \in \mathbb{Z} \quad \forall v \in \Lambda\}. \tag{6.61}$$

[4] This allows one to define the Néron-Severi group of every variety. It need not be torsion free.

It is the $\mathbb{Z}$-dual of Λ with respect to the real pairing between Ω and V. Then $\hat{\Lambda}$ is the kernel of the surjective map

$$\Omega \longrightarrow \mathrm{Pic}^0(X), \quad l \longmapsto e^{2\pi i \cdot \mathrm{Im} l}. \tag{6.62}$$

We get

$$\Omega/\hat{\Lambda} \cong \mathrm{Pic}^0(X). \tag{6.63}$$

Therefore $\mathrm{Pic}^0(X)$ has a canonical complex structure induced from that of Ω.

Before giving a theorem linking X to $X^\vee$, let us give a few definitions. Let $\mathcal{L}$ be a line bundle on X. Define

$$\phi_{\mathcal{L}}(x) = T_x^*\mathcal{L} \otimes \mathcal{L}^{-1}, \tag{6.64}$$

where

$$T_x : X \longrightarrow X, \quad T_x(y) = x + y, \tag{6.65}$$

is the translation-by-x morphism and T_x^* denotes the pullback of the line bundle $\mathcal{L}$ by T_x. Recall that $\mathcal{L}_{(H,\chi)}^{-1} = \mathcal{L}_{(-H,\chi^{-1})}$. One can show:

$$T_x^*\mathcal{L}_{(H,\chi)} \cong \mathcal{L}_{(H,\chi\mu_x)}, \tag{6.66}$$

with $\mu_x(\lambda) = e^{2\pi i E(x,\lambda)}$, $\lambda \in \Lambda$. Here, we extend E to V by $\mathbb{R}$-linearity.

Exercise 6.25. Prove the identity (6.66) by comparing the factors of automorphy of both sides.

Thus, $\phi_{\mathcal{L}} : X \longrightarrow \mathrm{Pic}^0(X)$ is given by a closed formula

$$\phi_{\mathcal{L}}(x) = e^{2\pi i (E(x,-))} \in \mathrm{Hom}(\Lambda, \mathbb{C}_1). \tag{6.67}$$

Corollary 6.26. *The map $\phi_{\mathcal{L}}$ is surjective if and only if E is non-degenerate.*

Definition 6.27. A *polarization* of X is a homomorphism $f : X \longrightarrow X^\vee$ such that $f = \phi_{\mathcal{L}}$ for some ample line bundle $\mathcal{L}$.

Remark 6.28.
1. A polarization is an *isogeny*, that is, it is a surjective homomorphism of abelian varieties with finite kernel. The *degree* of the polarization is the order of the kernel. It is its degree as a finite morphism.
2. A polarization is an isogeny, but the converse is not true. Indeed, if ϕ is a polarization then $-\phi$ is an isogeny that is never a polarization.

Exercise 6.29. Prove the last remark.

Definition 6.30. A polarization is *principal* if it is an isomorphism.

Exercise* 6.31. Find an abelian variety having no principal polarization.

Exercise 6.32. Let $\mathcal{L} = \mathcal{L}_{(H,\chi)}$ and $E = \mathrm{Im}(H)$. Prove that $\phi_{\mathcal{L}}$ is a principal polarization iff the pairing $E : \Lambda \times \Lambda \longrightarrow \mathbb{Z}$ is surjective.

We remark that for any abelian variety A, we can always find a principally polarized abelian variety B such that A is isogenous to B. This follows quite easily from the theory we have developed above and the reader may enjoy proving it to himself. In the case of arbitrary base field this follows from the finite Heisenberg group associated by Mumford to $(X, \mathcal{L})$.

One can parameterize abelian varieties, endowed with a rigid level structure and a polarization, by a fine moduli scheme. Let $\mathcal{H}_g$ be Siegel's upper half space

$$\left\{ \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \tau \in M_g(\mathbb{C}) : \tau = \tau^t, \mathrm{Im}(\tau) \gg 0 \right\}. \tag{6.68}$$

It turns out that the coarse moduli scheme of couples (A, λ), A a g-dimensional complex abelian variety, λ a principal polarization of A, is $\mathrm{Sp}(2g, \mathbb{Z})\backslash\mathcal{H}^g$. The action of $\mathrm{Sp}(2g, \mathbb{Z})$ is given by $\tau \mapsto (A\tau + B)(C\tau + D)^{-1}$.

THEOREM 6.33. *Let X be a complex torus, $X = V/\Lambda$. Let $X^\vee$ be the dual complex torus with its canonical complex structure.*

1. *$X^{\vee\vee}$ is canonically isomorphic to X.*
2. *$\mathrm{Hom}(X, X^\vee)$ consists of bilinear forms H on V such that $\mathrm{Im}H(\Lambda, \Lambda) \subseteq \mathbb{Z}$ ($\mathbb{C}$-linear in the first variable, and $\mathbb{C}$-anti-linear in the second variable).*
3. *The map $\phi \in \mathrm{Hom}(X, X^\vee)$ is of the form $\phi_{\mathcal{L}}$, for some $\mathcal{L}$, iff the associated form H is hermitian, i.e. $H(x, y) = \overline{H(y, x)}$. It is a polarization iff H is a positive definite Riemann form with respect to Λ.*
4. *Let $f : X \longrightarrow Y$ be a homomorphism of complex tori. Consider the following diagram:*

$$\begin{array}{ccc} \Omega(Y) & \xrightarrow{\widetilde{f^*}} & \Omega(X) \\ \downarrow & & \downarrow \\ \mathrm{Pic}^0(Y) & \xrightarrow[f^*]{} & \mathrm{Pic}^0(X) \end{array} \tag{6.69}$$

where $f^* = f^\vee$ is the pullback of line bundles, and $\widetilde{f^*}$ the lifting to the universal covering space. Then $\widetilde{f^*}$ is just the pull-back

$$f^* : \mathrm{Hom}_{\overline{\mathbb{C}}}(\Omega(Y), \mathbb{C}) \longrightarrow \mathrm{Hom}_{\overline{\mathbb{C}}}(\Omega(X), \mathbb{C}). \tag{6.70}$$

It follows easily that

$$(f + g)^\vee = f^\vee + g^\vee, \quad (fg)^\vee = g^\vee f^\vee, \quad f^{\vee\vee} = f. \tag{6.71}$$

5. *Let $\Phi : X \longrightarrow X^\vee$ be a polarization. The Weil pairing associated to Φ is defined as:*

$$\begin{aligned} X[n] \times X[n] &\xrightarrow{1\times\Phi} X[n] \times X^\vee[n] \\ &\xrightarrow{\cong} \tfrac{1}{n}\Lambda/\Lambda \times \tfrac{1}{n}\hat{\Lambda}/\hat{\Lambda} \\ &\longrightarrow \mu_n, \end{aligned} \tag{6.72}$$

via the map

$$\begin{aligned} \frac{1}{n}\Lambda/\Lambda \times \frac{1}{n}\hat{\Lambda}/\hat{\Lambda} &\longrightarrow \mu_n(\mathbb{C}) \\ (x, k) &\longmapsto e^{2\pi i \cdot n \cdot k(x)}, \end{aligned} \tag{6.73}$$

where we think of $\hat{\Lambda}$ as $\{k \in \Omega = \mathrm{Hom}_{\mathbb{R}}(V, \mathbb{R}) | k(\lambda) \in \mathbb{Z}\ \forall \lambda \in \Lambda\}$. The Weil pairing is an alternating bilinear pairing.

EXERCISE 6.34. Complete the proof of the above theorem. At this point it is mainly unfolding the definitions. Prove also that the Weil pairing is perfect iff $(\deg(\Phi), n) = 1$.

DEFINITION 6.35. Let B be a semi-simple algebra over $\mathbb{Q}$. Let Tr denote its reduced trace to $\mathbb{Q}$. By an *involution* on B we understand an isomorphism $\rho : B \longrightarrow B$, $x \mapsto x^\rho$ of $\mathbb{Q}$ vector spaces such that $(xy)^\rho = y^\rho x^\rho$. An element of B is called symmetric (with respect to ρ) if $x = x^\rho$. When ρ is understood we shall also denote it by $x \mapsto x^*$.

An involution $x \mapsto x^*$ on B is a *positive involution* if $\mathrm{Tr}(xx^*) > 0$ for all $x \neq 0$. Given a positive involution $x \mapsto x^*$ we say that an element $b \in B$ is positive if $\mathrm{Tr}(xbx^*) > 0$ for any $x \neq 0$.

DEFINITION 6.36. Let λ be a polarization on an abelian variety A, and put $\mathrm{End}^0(A) = \mathrm{End}(A) \otimes \mathbb{Q}$. The *Rosati involution* associated to λ is the map

$$\begin{aligned} \mathrm{End}^0(A) &\longrightarrow \mathrm{End}^0(A) \\ f &\mapsto f^* := \lambda^{-1} f^\vee \lambda \end{aligned} \tag{6.74}$$

FACT 6.37. The Rosati involution is a positive involution on $\mathrm{End}^0(A)$. See [**67**, Theorem 1.8, Chapter 5].

The semi-simple rational finite-dimensional algebras were classified by Albert. Since every abelian variety A has some polarization, it follows that the endomorphism algebras of abelian varieties all fall into this category. We follow [**105**, Section 1].

Every division algebra over $\mathbb{Q}$ with a positive involution belongs to the following four types of algebras.

1. (Type I) Totally real algebraic number field L.
2. (Type II) Central simple algebra B over L such that the simple components of $B \otimes \mathbb{R}$ are all isomorphic to $M_2(\mathbb{R})$.
3. (Type III) Central simple algebra B over L such that the simple components of $B \otimes \mathbb{R}$ are all isomorphic to the Hamilton quaternions over $\mathbb{R}$.
4. (Type IV) Central simple algebra B over a totally imaginary quadratic extension of L.

Recall that the canonical involution σ is defined by $x \mapsto \sigma(x) = x - \mathrm{Tr}(x)$. The positive involutions are respectively:

1. The identity.
2. Let $a \in B$ be an element such that a^2 is a totally negative element of L, we $\rho(x) = a\sigma(x)a^{-1}$. Then ρ is a positive involution and every positive involution ρ is obtained that way.
3. The canonical involution.
4. This case is more subtle. We only remark that if B is a CM field containing L then the involution is complex conjugation.

Let $\lambda : A \longrightarrow A^t$ be a polarization. Define a map

$$NS^0(A) \hookrightarrow \mathrm{End}^0(A), \tag{6.75}$$

$$\phi_\lambda : \mathcal{L}_{(H,\chi)} \mapsto \lambda^{-1}\phi_L. \tag{6.76}$$

FACT 6.38. The set $\phi_\lambda(NS^0(A))$ is composed of the symmetric elements of $\mathrm{End}^0(A)$ (under the Rosati involution), and the set $\phi_\lambda(NS^0(A)^+)$ is made of the positive symmetric elements of $\mathrm{End}^0(A)$. See [**67**, Chapter V, Proposition 2.1]

EXAMPLE 6.39. **Elliptic curves**
We finish this section by examining explicitely some of the results discussed above in the particular case of elliptic curves.

Let

$$\Lambda = \mathbb{Z} + \mathbb{Z}\tau, \quad \mathrm{Im}(\tau) > 0. \tag{6.77}$$

Let $\mathcal{E}$ be the elliptic curve $\mathbb{C}/\Lambda$. Let $E : \Lambda \times \Lambda \longrightarrow \mathbb{Z}$ be the real alternating pairing determined by $E(\tau, 1) = 1$. We have

$$E(a\tau + b, c\tau + d) = ad - bc. \tag{6.78}$$

How does one determine the associated Hermitian form H? If $H(1,1) = \alpha$, then $H(z_1, z_2) = z_1\overline{z_2}\alpha$. Recall that $H(x,y) = E(ix,y) + iE(x,y)$, so

$$\begin{aligned} \alpha = H(1,1) &= E(i,1) + iE(1,1) \\ &= E(i,1) \\ &= E\left(\frac{1}{\mathrm{Im}(\tau)} \cdot \tau - \frac{\Re(\tau)}{\mathrm{Im}(\tau)}, 1\right) \\ &= E\left(\frac{1}{\mathrm{Im}(\tau)} \cdot \tau, 1\right) = \frac{1}{\mathrm{Im}(\tau)} > 0, \end{aligned} \tag{6.79}$$

since $\tau \in \mathcal{H}$. Note also that $\mathrm{Im}(\tau)$ = area of a fundamental domain of the lattice Λ. Hence H is a positive definite Riemann form with respect to Λ. Since every other Hermitian form is proportional to H by a positive real scalar, such a form would be a Riemann form with respect to Λ if and only if this scalar is a positive integer, because the associated real form must be integer valued on $\Lambda \times \Lambda$. We deduce that H generates the Néron-Severi group of $\mathcal{E}$.

Let us calculate the dual lattice $\hat{\Lambda}$. Recall the procedure of defining the dual abelian variety. It is defined as a quotient of $\Omega = \mathrm{Hom}_{\overline{\mathbb{C}}}(\mathbb{C}, \mathbb{C})$ by a lattice dual to Λ. We use the identification

$$\mathbb{C} \longrightarrow \mathrm{Hom}_{\overline{\mathbb{C}}}(\mathbb{C}, \mathbb{C}), \quad c \longmapsto l_c, \tag{6.80}$$

where $l_c(1) = c$, and hence, $l_c(a + b\tau) = \overline{a + b\tau} \cdot c$. We write a general element in $\mathbb{C}$ as $\alpha + \beta\tau$, where α and β are real numbers, and obtain

(6.81)

$$\begin{aligned} \hat{\Lambda} &= \{c : \mathrm{Im}(l_c(\Lambda)) \subseteq \mathbb{Z}\} \\ &= \{c : \mathrm{Im}(l_c(1)), \mathrm{Im}(l_c(\tau)) \in \mathbb{Z}\} \\ &= \{c : \mathrm{Im}(c) \in \mathbb{Z}, \mathrm{Im}(\overline{\tau}c) \in \mathbb{Z}\} \\ &= \{\alpha + \beta\tau : \mathrm{Im}(\alpha + \beta\tau) = \beta \cdot \mathrm{Im}(\tau) \in \mathbb{Z},\ \mathrm{Im}(\alpha\overline{\tau} + \beta\tau\overline{\tau}) = -\alpha \cdot \mathrm{Im}(\tau) \in \mathbb{Z}\} \\ &= \frac{1}{\mathrm{Im}(\tau)}\Lambda \\ &\cong \Lambda \end{aligned}$$

Therefore the dual elliptic curve is isomorphic to the original.

Choose any semi-character χ with respect to H. We have the polarization

$$\mathcal{E} \xrightarrow{\Phi_{\mathcal{L}(H,\chi)}} \mathcal{E}^\vee. \tag{6.82}$$

under which $y \longmapsto e^{2\pi i E(y,-)}$. Let us calculate the Weil pairing:

$$\frac{1}{n}\Lambda \times \frac{1}{n}\Lambda \longrightarrow \frac{1}{n}\Lambda \times \frac{1}{n}\hat{\Lambda} \longrightarrow \mu_n. \tag{6.83}$$

By definition

$$(x, y) \mapsto (x, \Phi_{\mathcal{L}(H,\chi)}(y)) \mapsto (\Phi_{\mathcal{L}(H,\chi)}(y)(nx)), \tag{6.84}$$

and in particular,

$$(1/n, \tau/n) \mapsto (1/n, e^{2\pi i E(\tau/n,-)}) \mapsto e^{2\pi i E(\tau/n,1)} = e^{2\pi i/n}. \tag{6.85}$$

CHAPTER 2

Complex Abelian Varieties with Real Multiplication and Hilbert Modular Forms

1. Algebraic Preliminaries

Let L be a field extension of $\mathbb{Q}$.

DEFINITION 1.1. A field L is a *totally real field* of degree $[L:\mathbb{Q}] = g$, if every embedding $\sigma \in \mathrm{Emb}(L, \mathbb{C}) = \{\sigma_1, \ldots, \sigma_g\}$ factorizes via $\mathbb{R}$:

$$L \underset{\sigma_i}{\hookrightarrow} \mathbb{R}, \quad \forall i \in \{1, \ldots, g\}. \tag{1.1}$$

We will assume henceforth that L is totally real. Here are some important examples:

1. $L = \mathbb{Q}$;
2. $L = \mathbb{Q}(\sqrt{D})$, for $D > 1$ square-free;
3. $L = \mathbb{Q}(\zeta_m)^+$, i.e. the subfield of $\mathbb{Q}(\zeta_m)$ fixed under complex conjugation, where ζ_m is a primitive $m - th$ root of unity.

Given any set $S \subset L$ we denote by S^+ the *totally positive elements of* S. Namely, the elements $s \in S$ such that $\sigma_i(s) > 0$ for all i.

We let $\mathcal{O}_L$ denote the ring of integers of L. We let $\mathcal{D}_L = \mathcal{D}_{L/\mathbb{Q}}$ denote the different ideal of L over $\mathbb{Q}$. Thus,

$$\mathcal{D}_{L/\mathbb{Q}}^{-1} = \left\{\ell \in L : \mathrm{Tr}_{L/\mathbb{Q}}(\ell r) \in \mathbb{Z} \ \ \forall r \in \mathcal{O}_L\right\}. \tag{1.2}$$

We let $d_L = \mathrm{Norm}(\mathcal{D}_{L/\mathbb{Q}})$ denote the discriminant of L.

Let $Cl(L)$ stand for the ideal class group of L and let h denote its order. Let $Cl(L)^+$ stand for the strict ideal class group in L and denote its order by h^+. Thus, if $\mathfrak{A}, \mathfrak{B}$ are two fractional ideals of L then $\mathfrak{A} = \mathfrak{B}$ in $Cl(L)$ (resp. $Cl(L)^+$) if and only if there exists $\lambda \in L^\times$ (resp. $\lambda \in L^{\times +}$) such that $\mathfrak{A} = \lambda\mathfrak{B}$.

Note the exact sequence

$$1 \longrightarrow L^\times/{\mathcal{O}_L}^* L^{\times +} \longrightarrow Cl(L)^+ \longrightarrow Cl(L) \longrightarrow 1. \tag{1.3}$$

There is also an isomorphism $L^\times/L^{\times +} \longrightarrow \{\pm 1\}^g$ taking an element ℓ to the signs of its embeddings $(\mathrm{sign}(\sigma_1(\ell)), \cdots, \mathrm{sign}(\sigma_g(\ell)))$. Thus $[Cl(L)^+ : Cl(L)]$ divides 2^g.

EXERCISE 1.2. Show that $[Cl(L)^+ : Cl(L)] = [({\mathcal{O}_L}^\times)^+ : ({\mathcal{O}_L}^\times)^2]$.

There is an equivalent way of defining $Cl(L)^+$. Recall that $Cl(L)$ is the group of projective $\mathcal{O}_L$-module of rank 1 up to isomorphism. Indeed, every such module is isomorphic to a fractional ideal and that ideal is determined up to multiplication by $\lambda \in L^\times$. We define an *$\mathcal{O}_L$-module with a notion of positivity* to be a projective $\mathcal{O}_L$-module M of rank 1 together with a linear order $<_i$ on the real vector space $M \otimes_{\sigma_i} \mathbb{R}$ for every i. That is, a choice of connected component of $M \otimes_{\sigma_i} \mathbb{R}$.

We may then talk about the group of isomorphism classes of projective $\mathcal{O}_L$ - modules of rank 1 with a notion of positivity that we call temporarily $Cl(L)^{++}$. An isomorphism of $\mathcal{O}_L$ modules M_1, M_2 with notion of positivity, is an isomorphism of modules $\phi : M_1 \longrightarrow M_2$ such that the induced maps $\phi \otimes_{\sigma_i} 1$ respect the ordering.

One verifies that the natural map $Cl(L)^{++} \longrightarrow Cl(L)$ has fibers that are principal homogenous spaces under $L^\times/{\mathcal{O}_L}^* L^{\times +}$ and that there is a natural map

$$Cl(L)^{+} \longrightarrow Cl(L)^{++}, \tag{1.4}$$

obtained by taking on a fractional ideal $\mathfrak{A}$ the natural orderings induced from the embeddings $L \xrightarrow{\sigma_i} \mathbb{R}^g$, where $i = 1, \ldots, g$. It is easily checked that this map is injective and this identifies $Cl(L)^{+}$ with $Cl(L)^{++}$.

We also remark that two notions of positivity on M are equal if and only if they define the same positive cone. The positive cone consists of the elements of M that are ≥ 0 under any of the given orders on M.

2. Complex Abelian Varieties with Real Multiplication

Let L be a totally real field of degree g over $\mathbb{Q}$.

Definition 2.1. A complex abelian variety with *real multiplication* (abbreviated RM) by $\mathcal{O}_L$ is a g-dimensional abelian variety over $\mathbb{C}$ together with a *given* embedding

$$\iota : \mathcal{O}_L \hookrightarrow \mathrm{End}(A). \tag{2.1}$$

Example 2.2.

1. For $L = \mathbb{Q}$, elliptic curves satisfy the conditions of the definition.
2. Consider the product $E \times E$ and the totally real field $L = \mathbb{Q}(\sqrt{D})$ with $D > 1$ square free. The action of $\sqrt{D}$ (respectively $\frac{1+\sqrt{D}}{2}$) is given by a matrix of the form $\begin{pmatrix} a & b \\ c & -a \end{pmatrix}$ (resp. $\begin{pmatrix} (1+a)/2 & b/2 \\ c/2 & (1-a)/2 \end{pmatrix}$, and a is odd, b, c are even), when $D \equiv 2, 3 \mod 4$ (resp. $D \equiv 1 \mod 4$). Note that $D = a^2 + bc$.
3. For any elliptic curve E, say over $\mathbb{C}$, consider:

$$E \otimes_{\mathbb{Z}} \mathcal{O}_L \cong E^g \tag{2.2}$$

with the canonical right $\mathcal{O}_L$ action. The isomorphism being obtained by choosing a $\mathbb{Z}$-basis to $\mathcal{O}_L$. We get an abelian variety over $\mathbb{C}$ such that

$$(E \otimes_{\mathbb{Z}} \mathcal{O}_L)(\mathbb{C}) = E(\mathbb{C}) \otimes_{\mathbb{Z}} \mathcal{O}_L. \tag{2.3}$$

4. If A has RM by $\mathcal{O}_L$, so does the dual abelian variety $A^\vee$. That is, given a map

$$\iota : \mathcal{O}_L \longrightarrow \mathrm{End}(A), \tag{2.4}$$

we define

$$\iota^\vee : \mathcal{O}_L \longrightarrow \mathrm{End}(A^\vee) \tag{2.5}$$

by

$$\iota^\vee(\phi) = \phi^\vee, \tag{2.6}$$

where $\phi^\vee$ is the dual map to ϕ.

5. First, we say that an abelian variety is *simple* if it has no proper sub-abelian variety. For example, an elliptic curve is simple. By replacing the notion of isomorphism by the notion of isogeny the simple abelian varieties become the basic building blocks of the category of abelian variety. More precisely:

THEOREM 2.3. (Poincaré) *The category of abelian varieties considered up to isogeny is semi-simple with simple objects given by simple abelian varieties, i.e. every abelian variety is isogenous to a product*

$$A_1^{n_1} \times \cdots A_k^{n_k}, \quad A_i \text{ simple }, \tag{2.7}$$

and the n_i, A_i, are uniquely determined by A up to isogeny.

We say that A is *iso-simple* if A is isogenous to B^r for some simple abelian variety B. We shall see in Corollary 2.6 that every abelian with RM is iso-simple.

Let $J_0(p) = \mathrm{Jac}(X_0(p)) = \mathrm{Jac}(\Gamma_0(p)\backslash\mathcal{H}^*)$. Then it is a fact that the simple factors (the A_i's) of $J_0(p)$ are abelian varieties with real multiplication (by some field L_i).

2.1. Complex and rational representations.
The structure map,

$$\iota : \mathcal{O}_L \hookrightarrow \mathrm{End}(A), \tag{2.8}$$

can be extended by tensoring with $\mathbb{Q}$ to:

$$\iota \otimes 1 : \mathcal{O}_L \otimes \mathbb{Q} = L \hookrightarrow \mathrm{End}_{\mathbb{Q}}(A) = \mathrm{End}(A) \otimes \mathbb{Q}. \tag{2.9}$$

Suppose that $A = \mathbb{C}^g/\Lambda$. Any endomorphism r of A induces an endomorphism $\widetilde{r}$ on the universal covering space $\mathbb{C}^g$ such that the following diagram commutes:

$$\begin{array}{ccc} \mathbb{C}^g & \xrightarrow{\widetilde{r}} & \mathbb{C}^g \\ \downarrow & & \uparrow \\ A & \xrightarrow{r} & A \end{array} \tag{2.10}$$

We define the complex representation to be the map:

$$\rho_{\mathbb{C}} : \mathrm{End}_{\mathbb{Q}}(A) \longrightarrow \mathrm{GL}_g(\mathbb{C}), \quad r \mapsto \widetilde{r}. \tag{2.11}$$

We have an integral representation:

$$\rho_{\mathbb{Q}} : \mathrm{End}(A) \longrightarrow \mathrm{End}_{\mathbb{Z}}(\Lambda) \cong M_{2g}(\mathbb{Z}). \tag{2.12}$$

The last isomorphism requires a choice of a $\mathbb{Z}$-basis for Λ. It gives the rational representation as the map

$$\rho_{\mathbb{Q}} : \mathrm{End}_{\mathbb{Q}}(A) \longrightarrow \mathrm{End}_{\mathbb{Q}}(\Lambda \otimes \mathbb{Q}) \cong M_{2g}(\mathbb{Q}). \tag{2.13}$$

PROPOSITION 2.4. *The rational map $\rho_{\mathbb{Q}}$ is conjugate to $\rho_{\mathbb{C}} \oplus \overline{\rho_{\mathbb{C}}}$.*

PROOF. Let $N = \rho_{\mathbb{C}}(r)$ be the complex representation of an endomorphism $r \in \mathrm{End}_{\mathbb{Q}}(A)$: Put $\mathrm{Re}(N) = \frac{N+\overline{N}}{2}$, $\mathrm{Im}(N) = \frac{N-\overline{N}}{2i}$. We have the matrix

$$M = \begin{pmatrix} \mathrm{Re}(N) & -\mathrm{Im}(N) \\ \mathrm{Im}(N) & \mathrm{Re}(N) \end{pmatrix}. \tag{2.14}$$

The matrix M is describing the action of $N = \rho_{\mathbb{C}}(r)$ with respect to the basis

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}, \begin{pmatrix} i \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ i \end{pmatrix}. \tag{2.15}$$

Since $\Lambda \otimes \mathbb{R} = \mathbb{C}^g$, M is $\rho_{\mathbb{Q}}(r)$ up to conjugation. Thus, we only need to show that M is conjugate to $\begin{pmatrix} N & 0 \\ 0 & \overline{N} \end{pmatrix}$. But

$$\begin{pmatrix} I & iI \\ iI & I \end{pmatrix} \begin{pmatrix} \mathrm{Re}N & -\mathrm{Im}(N) \\ \mathrm{Im}(N) & \mathrm{Re}(N) \end{pmatrix} \begin{pmatrix} \frac{1}{2} \end{pmatrix} \begin{pmatrix} I & -iI \\ -iI & I \end{pmatrix} = \begin{pmatrix} N & 0 \\ 0 & \overline{N} \end{pmatrix}. \tag{2.16}$$

□

COROLLARY 2.5. *The tangent space $\mathfrak{t}_{A,0}$ to A at the origin is a free $\mathcal{O}_L \otimes \mathbb{C}$-module of rank 1.* [1]

PROOF. Note that the tangent space $\mathfrak{t}_{A,0}$ to A at the origin is canonically $\mathbb{C}^g$. Thus it is left to check that each σ_i appears exactly (equivalently, at least) once in the action of L on $\mathfrak{t}_{A,0}$. This action is described by the complex representation of L as a subspace of $\mathrm{End}_{\mathbb{Q}}(A)$. Since L is totally real the complex conjugate of the complex representation of L is equivalent to itself, and hence, by Proposition 2.4, the rational representation is equal to twice the complex representation. However, in the rational representation (being rational) every character σ_i has to appear with the same multiplicity. Hence, every σ_i appears in the complex representation as well. □

COROLLARY 2.6. *Any complex abelian variety with RM is iso-simple.*[2]

PROOF. The follows from the following considerations. Let B^r be a maximal iso-simple factor of A. Then $L \hookrightarrow \mathrm{GL}_r(\mathrm{End}(B))$ and the tangent space to B at the origin is a free $\mathcal{O}_L \otimes \mathbb{C}$-module. Hence $g | \dim(B^r)$. □

2.2. Construction of families of abelian varieties with real multiplication.

Let us fix some notation. For $\ell \in L$ and $t = (t_1, \ldots, t_g) \in \mathbb{C}^g$ we put

$$\ell \cdot t = \ell \cdot (t_1, \ldots, t_g) = (\sigma_1(\ell)t_1, \ldots, \sigma_g(\ell)t_g). \tag{2.17}$$

For the rest of the chapter, fix fractional ideals $\mathfrak{A}, \mathfrak{B}$ of $\mathcal{O}_L$.

Given $z = (z_1, \ldots, z_g) \in \mathcal{H}^g$, put

$$\begin{aligned} \Lambda_z &= \mathfrak{A} \cdot (z_1, \ldots, z_g) + \mathfrak{B} \cdot (1, \ldots, 1) \\ &= \{(\sigma_1(a)z_1 + \sigma_1(b), \ldots, \sigma_g(a)z_g + \sigma_g(b)) : a \in \mathfrak{A}, b \in \mathfrak{B}\} \subseteq \mathbb{C}^g. \end{aligned} \tag{2.18}$$

Note that Λ_z is a lattice: First,

$$\text{rank } _{\mathbb{Z}}(\Lambda_z) \leq \text{ rank } _{\mathbb{Z}}\mathfrak{A} + \text{ rank } _{\mathbb{Z}}\mathfrak{B} \leq 2g. \tag{2.19}$$

On the other hand,

$$\Lambda_z \otimes_{\mathbb{Z}} \mathbb{R} \cong \prod_{i=1}^{g} (\mathbb{R}z_i + \mathbb{R}) \cong \mathbb{C}^g, \tag{2.20}$$

[1] In the context of this Corollary, see Chapter 3, Section 5.

[2] In fact, this holds for every abelian variety with real multiplication

since $\mathrm{Im}(z_i) > 0$. Thus Λ_z must be of rank $2g$ and contains an $\mathbb{R}$-basis to $\mathbb{C}^g$. That is, Λ_z is a lattice.

We let A_z be the complex torus $\mathbb{C}^g/\Lambda_z$. We now proceed to construct polarizations on A_z. Using the connection between Riemann forms and polarizations (see Section 6.1), we turn our attention to constructing certain $\mathbb{R}$-bilinear antisymmetric forms.

Let $r \in L$ and define

$$E_r : \mathfrak{A} \oplus \mathfrak{B} \times \mathfrak{A} \oplus \mathfrak{B} \longrightarrow \mathbb{Q}, \tag{2.21}$$

$$E_r((x_1, y_1), (x_2, y_2)) = \mathrm{Tr}_{L/\mathbb{Q}}(r(x_1 y_2 - x_2 y_1)). \tag{2.22}$$

The proof of the following Lemma is left to the reader

LEMMA 2.7. 1. *E_r is an alternating bilinear form.*
2. *The image of E_r is in $\mathbb{Z}$ if and only if $r \in (\mathcal{D}_{L/\mathbb{Q}}\mathfrak{A}\mathfrak{B})^{-1}$.*

Note that for all $\ell \in \mathcal{O}_L$, $\alpha \in \mathfrak{A} \oplus \mathfrak{B}, \beta \in \mathfrak{A} \oplus \mathfrak{B}$,

$$E_r(\ell\alpha, \beta) = E_r(\alpha, \ell\beta). \tag{2.23}$$

Transport E_r to Λ_z and extend $\mathbb{R}$-linearly to $\mathbb{C}^g$. We get an antisymmetric $\mathbb{R}$-bilinear pairing

$$E_{r,z} : \mathbb{C}^g \times \mathbb{C}^g \longrightarrow \mathbb{R}, \tag{2.24}$$

with the property

$$E_{r,z}(\ell \cdot t, t') = E_{r,z}(t, \ell \cdot t'). \tag{2.25}$$

Moreover

$$E_{r,z}(\Lambda_z, \Lambda_z) \subseteq \mathbb{Z} \Leftrightarrow r \in (\mathcal{D}_{L/\mathbb{Q}}\mathfrak{A}\mathfrak{B})^{-1}. \tag{2.26}$$

If $r \neq 0$, $E_{r,z}$ is also perfect, by property of the trace. It also follows from the following:

LEMMA 2.8. *Let*

$$H_{r,z}((x_1, \ldots, x_g), (y_1, \ldots, y_g)) = \sum_{i=1}^{g} \frac{x_i \overline{y_i} \sigma_i(r)}{\mathrm{Im}(z_i)}. \tag{2.27}$$

Then $H_{r,z}$ is an hermitian form and $\mathrm{Im}H_{r,z} = E_{r,z}$.

PROOF. This form is clearly hermitian. To prove the last statement, it is enough to show

$$\mathrm{Im}H_{r,z}((x_1, \ldots, x_g), (y_1, \ldots, y_g)) = E_{r,z}((x_1, \ldots, x_g), (y_1, \ldots, y_g)), \tag{2.28}$$

for $(x_1, \ldots, x_g) \in \mathfrak{A} \cdot (z_1, \ldots, z_g)$ and $(y_1, \ldots, y_g) \in \mathfrak{B} \cdot (1, \ldots, 1)$. (Because $E_{r,z}$ and $\mathrm{Im}H_{r,z}$ are antisymmetric bilinear forms, they are determined by their values on such couples). But

$$\begin{aligned}
&\operatorname{Im}\left\{H\left(\sigma_1(a)z_1,\ldots,\sigma_g(a)z_g\right),\left(\sigma_1(b),\ldots,\sigma_g(b)\right)\right\}\\
&\qquad=\sum_{i=1}^{g}\operatorname{Im}\left\{\frac{\sigma_i(a)z_i\overline{\sigma_i(b)}\sigma_i(r)}{\operatorname{Im}(z_i)}\right\}\\
&\qquad=\sum_{i=1}^{g}\sigma_i(a)\sigma_i(b)\sigma_i(r) \qquad (2.29)\\
&\qquad=\operatorname{Tr}_{L/\mathbb{Q}}(abr)\\
&\qquad=E_r((a,0),(0,b))\\
&\qquad=E_{r,z}(a\cdot(z_1,\ldots,z_g),b\cdot(1,\ldots,1)).
\end{aligned}$$

□

We note that if $(x_2,y_2)\in\mathfrak{A}\oplus\mathfrak{B}$, then the set

$$\{(x_1,y_1)\in L\oplus L : E_r((x_1,y_1),(x_2,y_2))\in\mathbb{Z},\ \forall(x_2,y_2)\in\mathfrak{A}\oplus\mathfrak{B}\} \tag{2.30}$$

is precisely

$$(r\mathcal{D}_{L/\mathbb{Q}}\mathfrak{B})^{-1}\oplus(r\mathcal{D}_{L/\mathbb{Q}}\mathfrak{A})^{-1}. \tag{2.31}$$

Hence in the case $E_{r,z}$ defines a polarization, its degree is equal to

$$\begin{aligned}
&=[(r\mathcal{D}_{L/\mathbb{Q}}\mathfrak{B})^{-1}:\mathfrak{A}]\cdot[(r\mathcal{D}_{L/\mathbb{Q}}\mathfrak{A})^{-1}:\mathfrak{B}]\\
&=[(r\mathcal{D}_{L/\mathbb{Q}}\mathfrak{A}\mathfrak{B})^{-1}:\mathcal{O}_L]^2 \qquad (2.32)\\
&=\operatorname{Norm}(r\mathcal{D}_{L/\mathbb{Q}}\mathfrak{A}\mathfrak{B})^2.
\end{aligned}$$

Summing up, we get:

Corollary 2.9. *Let $z\in\mathcal{H}^g$ and $r\in L$. The form $E_{r,z}$ defines a polarization on A_z iff $r\in\left((\mathcal{D}_{L/\mathbb{Q}}\mathfrak{A}\mathfrak{B})^{-1}\right)^+$. If this holds, then A_z is an abelian variety with RM by $\mathcal{O}_L$, and the degree of the polarization given by $E_{r,z}$ is $\operatorname{Norm}(r\mathcal{D}_{L/\mathbb{Q}}\mathfrak{A}\mathfrak{B})^2$. In particular, there exists an r such that $E_{r,z}$ is principal iff $\mathfrak{A}\mathfrak{B}=\mathcal{D}_{L/\mathbb{Q}}^{-1}$ in $Cl(L)^+$.*

Definition 2.10. Let (A,ι) be an abelian variety with RM by $\mathcal{O}_L$. Let

$$\mathcal{M}_A:=\{\lambda:A\longrightarrow A^\vee:\lambda=\lambda^\vee,\ \lambda\text{ is an }\mathcal{O}_L\text{-linear homomorphism}\}\subseteq\operatorname{NS}(A), \tag{2.33}$$

$$\mathcal{M}_A^+:=\{\lambda\in\mathcal{M}_A:\lambda\text{ is a polarization}\}\subseteq\operatorname{NS}(A)^+. \tag{2.34}$$

We call $\mathcal{M}_A$ the *polarization module* of A and $\mathcal{M}_A^+$ the positive cone of polarizations.

Exercise 2.11. Let $\lambda:A\longrightarrow B$ be an isogeny of two (g-dimensional) abelian varieties with real multiplication. Show that if λ is $\mathcal{O}_L$-linear and non-zero then λ is an isogeny.

Lemma 2.12. 1. *$\mathcal{M}_A$ is a projective $\mathcal{O}_L$-module of rank 1. In particular,*

$$\ell\in\mathcal{O}_L,\lambda\in\mathcal{M}_A\implies\lambda\circ\ell\in\mathcal{M}_A. \tag{2.35}$$

2. *The set $\mathcal{M}_A^+$ is a positive cone. Moreover, $\mathcal{M}_A$ is generated by $\mathcal{M}_A^+$.*

PROOF. We first check that $\mathcal{M}_A$ is an $\mathcal{O}_L$-module. We write more pedantically $\lambda \circ \ell$ as $\lambda \circ \iota(\ell)$. Then $(\lambda \circ \iota(\ell))^\vee = \iota(\ell)^\vee \circ \lambda^\vee = \iota^\vee(\ell) \circ \lambda = \lambda \circ \iota(\ell)$, using that $\lambda^\vee = \lambda$ and that λ is $\mathcal{O}_L$-linear. It is also clear that $\lambda \circ \iota(\ell)$ is $\mathcal{O}_L$-linear. Thus $\mathcal{M}_A$ is an $\mathcal{O}_L$-module, torsion free (hence projective) by Exercise 2.11.

Next, we remark that any abelian variety A over $\mathbb{C}$ with RM by $\mathcal{O}_L$ has an $\mathcal{O}_L$ linear polarization. This follows from considerations as in Theorem 2.16. See also See [**97**, Proposition 1.17]. Thus $\mathcal{M}_A^+$ is not empty.

To see $\mathcal{M}_A$ is of rank one, take some $\lambda \in \mathcal{M}_A^+$ and get an embedding

$$\mathcal{M}_A \hookrightarrow \mathrm{Cent}_{\mathrm{End}^0(A)}(L)^{sym}, \tag{2.36}$$

where *sym* on the right hand side signifies the elements fixed by the Rosati involution defined by λ. The classification of endomorphism algebras with positive involution shows that we must have the image of $\mathcal{M}_A \otimes \mathbb{Q}$ equal to L. See Chapter 1, Section 6.2.

The rest follows from the general principles given there. The Rosati involution induced by some element of $\mathcal{M}_A^+$ must induce the identity on L. Thus the embedding $\mathcal{M}_A \hookrightarrow L$ given by (2.36) identifies $\mathcal{M}_A$ with a fractional ideal $\mathfrak{a}$ of L, and identifies $\mathcal{M}_A^+$ with $\mathfrak{a}^+$. □

In particular, $(\mathcal{M}_A, \mathcal{M}_A^+)$ is naturally an $\mathcal{O}_L$-module with a notion of positivity (via the embedding in $\mathrm{End}^0(A)$).

In contrast to the case of elliptic curves, $E = \mathbb{C}/\Lambda$, where all lattices Λ are indistinguishable as $\mathbb{Z}$-modules, in the case of a totally real field L of degree $g > 1$ and abelian varieties with RM, $\mathbb{C}^g/\Lambda$, the lattice Λ is a projective rank 2 module over $\mathcal{O}_L$ and as such has a canonically associated ideal class of $Cl(L)$, called its Steinitz class, See [**34**, Chapter II.4]. If we write $\Lambda \cong \mathfrak{A} \oplus \mathfrak{B}$, and in fact every projective rank 2 module can be written this way, then the Steinitz class is $\mathfrak{AB}$. More canonically, it is $\bigwedge^2_{\mathcal{O}_L} \Lambda$.

Those Steinitz classes give a discrete invariant of abelian varieties with RM by $\mathcal{O}_L$ and already show that components of the moduli space of abelian varieties with RM, map into $Cl(L)$. The consideration of polarization would show that there is a natural bijection between the components of the moduli space and $Cl(L)^+$.

DEFINITION 2.13. The group $\mathrm{GL}(\mathfrak{A} \oplus \mathfrak{B})^+$ consists of the matrices

$$\left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a, d \in \mathcal{O}_L, b \in \mathfrak{A}^{-1}\mathfrak{B}, c \in \mathfrak{AB}^{-1}, ad - bc \in (\mathcal{O}_L^\times)^+ \right\} \tag{2.37}$$

with matrix multiplication.

REMARK 2.14. Note that $\mathrm{GL}(\mathfrak{A} \oplus \mathfrak{B})^+$ is the set of matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ in $\mathrm{GL}_2(L)$ with totally positive determinant, such that

$$(\mathfrak{A}, \mathfrak{B}) \begin{pmatrix} a & b \\ c & d \end{pmatrix} = (\mathfrak{A}, \mathfrak{B})\,. \tag{2.38}$$

The group $G = \mathrm{GL}(\mathfrak{A} \oplus \mathfrak{B})^+$ acts on $\mathcal{H}^g$ from the left by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} z = \left(\dots, \frac{\sigma_i(a) z_i + \sigma_i(b)}{\sigma_i(c) z_i + \sigma_i(d)}, \dots \right). \tag{2.39}$$

We also define the special linear subgroup of $\mathrm{GL}(\mathfrak{A} \oplus \mathfrak{B})^+$:

DEFINITION 2.15. The group $\mathrm{SL}(\mathfrak{A}\oplus\mathfrak{B})$ is the subgroup of $\mathrm{GL}(\mathfrak{A}\oplus\mathfrak{B})^+$ composed of matrices with determinant 1.

THEOREM 2.16. 1. *The isomorphism classes of* $(A,\iota)/\mathbb{C}$ *such that there exists an isomorphism*

$$(\mathcal{M}_A,\mathcal{M}_A^+) \xrightarrow{\sim} (\mathfrak{C},\mathfrak{C}^+),\quad \mathfrak{C}=(\mathcal{D}_L\mathfrak{A}\mathfrak{B})^{-1}, \tag{2.40}$$

are parameterized by $\mathrm{GL}(\mathfrak{A}\oplus\mathfrak{B})^+\backslash\mathcal{H}^g$.

2. *The isomorphism classes of* $(A,\iota)/\mathbb{C}$ *together with a* given *isomorphism*

$$m:(\mathcal{M}_A,\mathcal{M}_A^+) \xrightarrow{\sim} (\mathfrak{C},\mathfrak{C}^+),\quad \mathfrak{C}=(\mathcal{D}_L\mathfrak{A}\mathfrak{B})^{-1} \tag{2.41}$$

are parameterized by $\mathrm{SL}(\mathfrak{A}\oplus\mathfrak{B})\backslash\mathcal{H}^g$.

REMARK 2.17. There is of course a difference between the two statements. There is more elbow room in the first case since any $x_\epsilon\in(\mathcal{O}_L^\times)^+$ induces an isomorphism:

$$(\mathfrak{C},\mathfrak{C}^+) \xrightarrow{\sim} (\mathfrak{C},\mathfrak{C}^+). \tag{2.42}$$

The possible m's form in fact a principal homogeneous space under $(\mathcal{O}_L^\times)^+$. Fix (A,ι,m). For any isomorphism $\phi:(B,j)\xrightarrow{\sim}(A,\iota)$ the diagram:

$$\begin{array}{ccc} B & \xrightarrow{\phi^*f} & B^\vee \\ \downarrow{\scriptstyle\phi} & & \uparrow{\scriptstyle\phi^\vee} \\ A & \xrightarrow{f} & A^\vee \end{array} \tag{2.43}$$

identifies $(\mathcal{M}_A,\mathcal{M}_A^+)$ with $(\mathcal{M}_B,\mathcal{M}_B^+)$. Let $m:(\mathcal{M}_A,\mathcal{M}_A^+)\cong(\mathfrak{C},\mathfrak{C}^+)$ be an isomorphism. Then the induced isomorphism $(\mathcal{M}_B,\mathcal{M}_B^+)\cong(\mathfrak{C},\mathfrak{C}^+)$ is the following: If $g\in\mathcal{M}_B$ then

$$\phi^*m(g)=m((\phi^{-1})^\vee g\phi^{-1}). \tag{2.44}$$

Now, take in particular $(B,j)=(A,\iota)$ and ϕ equal to multiplication by ϵ^{-1}, where $\epsilon\in\mathcal{O}_L^\times$. Then

$$\phi^*m(g)=m((\epsilon)^\vee g\epsilon)=m(g\epsilon^2). \tag{2.45}$$

Thus

$$(A,\iota,m)\cong(A,\iota,m\epsilon^2). \tag{2.46}$$

Therefore, for every $\mu\in(\mathcal{O}_L^\times)^2$,

$$(A,\iota,m)\cong(A,\iota,m\mu). \tag{2.47}$$

Generically, the only endomorphisms of (A,ι) are $\iota(\mathcal{O}_L)$, as one may verify directly taking $(A,\iota)=A_z$, where $z=(z_1,\dots,z_g)$ are independent variables over L.

Thus, generically, the map

$$(A,\iota,m)\longrightarrow(A,\iota,\exists m) \tag{2.48}$$

has degree $\left[(\mathcal{O}_L^\times)^+:(\mathcal{O}_L^\times)^2\right]=[\mathrm{PGL}(\mathfrak{A}\oplus\mathfrak{B})^+:\mathrm{PSL}(\mathfrak{A}\oplus\mathfrak{B})]$.

PROOF. (Of Theorem) Let $\mu = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ be an element of $\mathrm{GL}(\mathfrak{A} \oplus \mathfrak{B})^+$, and let

$$z = (z_1, \dots, z_g) \in \mathcal{H}^g, \;\; 1 = (1, \dots, 1). \tag{2.49}$$

We have

$$\Lambda_z = \mathfrak{A} \cdot z + \mathfrak{B} \cdot 1 \quad \Lambda_{\mu z} = \mathfrak{A} \cdot \mu z + \mathfrak{B} \cdot 1. \tag{2.50}$$

First, by direct calculation, we check that $(\mathfrak{A}, \mathfrak{B})\mu = (\mathfrak{A}, \mathfrak{B})$, for $\mu \in \mathrm{GL}(\mathfrak{A} \oplus \mathfrak{B})^+$. Consider the map

$$f : \mathbb{C}^g \longrightarrow \mathbb{C}^g, \;\; x \mapsto x \cdot M, \tag{2.51}$$

where

$$M = \mathrm{diag}(\sigma_1(c)z_1 + \sigma_1(d), \dots, \sigma_g(c)z_g + \sigma_g(d)). \tag{2.52}$$

Under this map

$$f(\Lambda_{\mu z}) = \Lambda_{\mu z}\mathrm{diag}(\dots, \sigma_i(c)z_i + \sigma_i(d), \dots) \tag{2.53}$$

is the lattice

$$\{\left(\dots, \sigma_i(\alpha)\left(\frac{\sigma_i(a)z_i + \sigma_i(b)}{\sigma_i(c)z_i + \sigma_i(d)}\right) + \sigma_i(\beta), \dots\right)\mathrm{diag}(\dots, \sigma_i(c)z_i + \sigma_i(d), \dots) : (\alpha, \beta) \in \mathfrak{A} \oplus \mathfrak{B}\}. \tag{2.54}$$

This lattice is equal to

$$\left\{\left(\dots, \sigma_i\left((\alpha, \beta)\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right)\begin{pmatrix} z_i \\ 1 \end{pmatrix}, \dots\right) : (\alpha, \beta) \in \mathfrak{A} \oplus \mathfrak{B}\right\}. \tag{2.55}$$

Because $(\mathfrak{A}, \mathfrak{B})\mu = (\mathfrak{A}, \mathfrak{B})$, for $\mu \in \mathrm{GL}(\mathfrak{A} \oplus \mathfrak{B})^+$, this set is precisely Λ_z. Thus f induces a map

$$f : A_{\mu z} \longrightarrow A_z. \tag{2.56}$$

This map an isomorphism of abelian varieties with $\mathcal{O}_L$-action, because L acts diagonally: for $\ell \in L$

$$\ell \longmapsto \begin{pmatrix} \sigma_1(\ell) & & \\ & \ddots & \\ & & \sigma_g(\ell) \end{pmatrix}, \tag{2.57}$$

because the action of ℓ on any element in $\mathfrak{A} \cdot z + \mathfrak{B} \cdot 1$ is

$$\alpha \cdot z + \beta \cdot \mapsto \ell\alpha \cdot z + \ell\beta \cdot 1 = (\alpha \cdot z + \beta \cdot 1)\begin{pmatrix} \sigma_1(\ell) & & \\ & \ddots & \\ & & \sigma_g(\ell) \end{pmatrix}. \tag{2.58}$$

This holds also for the extends action of L to $\mathbb{R} \otimes_{\mathbb{Z}} (\mathfrak{A} \cdot z + \beta \cdot 1) = \mathbb{C}^g$.

Recall that by Lemma 2.8

$$H_{r,z}((x_1, \dots, x_g), (y_1, \dots, y_g)) = (x_1, \dots, x_g)\mathrm{diag}\left(\dots, \frac{\sigma_i(r)}{\mathrm{Im}(z_i)}, \dots\right)^t (\overline{y_1}, \dots, \overline{y_g}). \tag{2.59}$$

Let us calculate $f^*H_{r,z}(x,y)$. We have

(2.60)

$$\begin{aligned} f^*H_{r,z}(x,y) = (x_1,\dots,x_g)\mathrm{diag}(\dots,\sigma_i(c)z_i+\sigma_i(d),\dots)\mathrm{diag}(\dots,\frac{\sigma_i(r)}{\mathrm{Im}(z_i)},\dots) \\ \mathrm{diag}(\dots,\sigma_i(c)\overline{z_i}+\sigma_i(d),\dots)^t(\overline{y_1},\dots,\overline{y_g}) \\ = (x_1,\dots,x_g)\mathrm{diag}(\dots,b_i,\dots)^t(\overline{y_1},\dots,\overline{y_g}), \end{aligned}$$

where

$$b_i = \sigma_i(r)\cdot\frac{|\sigma_i(c)z_i+\sigma_i(d)|^2}{\mathrm{Im}(z_i)} = \frac{\sigma_i(r)}{\mathrm{Im}(\mu z_i)}\sigma_i(ad-bc) = \frac{\sigma_i(r)}{\mathrm{Im}(\mu z)_i}\sigma_i(ad-bc). \tag{2.61}$$

Hence

$$f^*H_{r,z} = H_{(\det\mu)r,\mu z}. \tag{2.62}$$

Thus, if we want to get an isomorphism of (A_z,ι,m_z) to $(A_{\mu z},\iota,m_{\mu z})$, we need $\det\mu = 1$.

To prove the first statement of the theorem, we still need to prove the assertions **(A)** and **(B)** below.

(A) If $A_z \cong A_{z'}$ as abelian varieties with RM by $\mathcal{O}_L$, and $z,z'\in\mathcal{H}^g$, show that there exists a μ in $\mathrm{GL}(\mathfrak{A}\oplus\mathfrak{B})^+$ such that $\mu z = z'$.

Suppose that $A_z \cong A_{z'}$ as abelian varieties with RM by $\mathcal{O}_L$. Let $M:\mathbb{C}^g\longrightarrow\mathbb{C}^g$ be the transformation inducing the isomorphism:

$$A_{z'} = \mathbb{C}^g/\Lambda_{z'} \longrightarrow \mathbb{C}^g/\Lambda_z = A_z. \tag{2.63}$$

The action of L on $\mathfrak{t}_{A_z,0} = \mathfrak{t}_{A_{z'},0} = \mathbb{C}^g$ is diagonal, and since M commutes with this action and the σ_i are independent, M must be a diagonal matrix $\mathrm{diag}(m_1,\dots,m_g)$. We can write $m_i = \sigma_i(c)z_i+\sigma_i(d)$, for some $c,d\in L$. Then M takes $\mathfrak{B}\cdot 1\subseteq\Lambda_{z'}$ into Λ_z. That is, for every $\beta\in\mathfrak{B}$

$$((\sigma_1(c)z_1+\sigma_1(d))\sigma_1(\beta),\dots,(\sigma_g(c)z_g+\sigma_g(d))\sigma_g(\beta))\in\Lambda_z. \tag{2.64}$$

We must therefore have $c\in\mathfrak{B}^{-1}\mathfrak{A}, d\in\mathcal{O}_L$. On the other hand, any $\alpha\cdot z'$ (with $\alpha\in L$) is mapped into $\mathbb{Q}\otimes\Lambda_z$ .That is, for suitable $a,b\in L$:

$$\begin{aligned} ((\sigma_1(c)z_1+\sigma(d))\sigma_1(\alpha)\cdot z_1',\dots,(\sigma_g(c)z_g+\sigma_g(d))\sigma_g(\alpha)\cdot z_g') \\ = ((\sigma_1(a)z_1+\sigma_1(b),\dots,\sigma_1(a)z_g+\sigma_b(b)). \end{aligned} \tag{2.65}$$

Take $\alpha=1$, so $z_i' = \frac{\sigma_1(a)z_1+\sigma_1(b)}{\sigma_1(c)z_1+\sigma_1(d)}$, for $\begin{pmatrix} a & b\\ c & d\end{pmatrix} = \mu\in\mathrm{GL}_2(L)$. Since μ preserves lattices, it is actually in $\mathrm{GL}(\mathfrak{A}\oplus\mathfrak{B})$, and $\mu\cdot z = z'$ gives $\sigma_i(\det\mu)>0$ for all i. That is, $\mu\in\mathrm{GL}(\mathfrak{A}\oplus\mathfrak{B})^+$.

(B) Given (A,ι) such that $(\mathcal{M}_A,\mathcal{M}_A^+)\cong(\mathfrak{C},\mathfrak{C}^+)$, then $(A,\iota)\cong(A_z,\iota)$, for some $z\in\mathcal{H}^g$.

Write $A=\mathbb{C}^g/\Lambda$. Since Λ is a projective rank 2 $\mathcal{O}_L$-module, we may write $\Lambda\cong\mathfrak{A}'\oplus\mathfrak{B}'$ for some fractional $\mathcal{O}_L$-ideals $\mathfrak{A}',\mathfrak{B}'$, and the isomorphism class of Λ is determined by its Steinitz class $\mathfrak{A}'\mathfrak{B}'$ [**34**, Theorem 13, p.95]. See also 2.2. The

polarization module of $\mathbb{C}^g/\Lambda$ is $(\bigwedge^2_{\mathcal{O}_L}\Lambda)^* = (\mathcal{D}_{L/\mathbb{Q}}\mathfrak{A}'\mathfrak{B}')^{-1}$. Hence, $\Lambda \cong \mathfrak{A}\oplus\mathfrak{B}$. We know that $\rho_{\mathbb{C}} \sim \mathrm{diag}(\sigma_1(l),\dots,\sigma_g(l))$.

We may therefore choose coordinates on $\mathbb{C}^g$ such that: $A = \mathbb{C}^g/\Lambda, \Lambda \cong \mathfrak{A}\oplus\mathfrak{B}$ and the action of L is given by the diagonal map. If necessary, we may still change coordinates by diagonal matrices, so that the map:

$$\phi : \mathfrak{A}\oplus\mathfrak{B} \longrightarrow \Lambda, \tag{2.66}$$

extends to a linear map on:

$$\phi : L\oplus L \longrightarrow \mathbb{Q}\cdot\Lambda, \tag{2.67}$$

satisfying $\phi(0,1) = (1,\dots,1)$ and $\phi(1,0) = (z_1,\dots,z_g)$. Every L-linear Riemann form is easily seen to be of the form:

$$H_{(r,z)}(x,y) = \sum \frac{x_i\overline{y_i}\sigma_i(r)}{\mathrm{Im}z_i}, \quad r\in(\mathcal{D}_{L/\mathbb{Q}}\mathfrak{A}\mathfrak{B})^{-1}. \tag{2.68}$$

It defines a polarization, i.e., is positive definite, if and only if for every i we have $\mathrm{sign}(\sigma_i(r)) = \mathrm{sign}(\mathrm{Im}(z_i))$. We conclude that if we choose $\mathfrak{A},\mathfrak{B}$ such that $\Lambda\cong\mathfrak{A}\oplus\mathfrak{B}$ and $\mathrm{NS}(A)^+ \cong (\mathcal{D}\mathfrak{A}\mathfrak{B})^{-1+}$, then $\mathrm{Im}(z_i)>0$ for all i, hence $A = A_z$ for some $z\in\mathcal{H}^g$. □

Fix representatives of $Cl(L)^+$ of the form $\{(\mathfrak{A},\mathfrak{A}^+)\}$.

COROLLARY 2.18. 1.

$$\{\textit{Isomorphism classes of } (A,\iota)/\mathbb{C}\} \longleftrightarrow \coprod_{(\mathfrak{A},\mathfrak{A}^+)} \mathrm{GL}(\mathcal{O}_L\oplus\mathfrak{A})^+\backslash\mathcal{H}^g, \tag{2.69}$$

with (A,ι) parameterized by $\mathrm{GL}(\mathcal{O}_L\oplus\mathfrak{A})^+$ iff there exists an isomorphism

$$(\mathcal{M}_A,\mathcal{M}_A^+) \cong \left((\mathcal{D}_L\mathfrak{A}),(\mathcal{D}_L\mathfrak{A})^{-1\ +}\right). \tag{2.70}$$

2.

$$\{\textit{Isomorphism classes of } (A,\iota,m)/\mathbb{C}\} \longleftrightarrow \coprod_{(\mathfrak{A},\mathfrak{A}^+)} \mathrm{SL}(\mathcal{O}_L\oplus\mathfrak{A})\backslash\mathcal{H}^g. \tag{2.71}$$

REMARK 2.19. There is a way to compactify $\mathrm{GL}(\mathfrak{A}\oplus\mathfrak{B})^+\backslash\mathcal{H}^g$ that follows the lines of the classical theory for elliptic curves: We add the finite set of points $\mathrm{GL}(\mathfrak{A}\oplus\mathfrak{B})^+\backslash\mathbb{P}^1(L)$ to $\mathrm{GL}(\mathfrak{A}\oplus\mathfrak{B})^+\backslash\mathcal{H}^g$. The action of $\mathrm{GL}(\mathfrak{A}\oplus\mathfrak{B})^+$ on $\mathbb{P}^1(L)$ is the usual one:

$$\begin{pmatrix} a & b\\ c & d\end{pmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = \begin{pmatrix} a\alpha+b\beta\\ c\alpha+d\beta\end{pmatrix}. \tag{2.72}$$

There is charm in the following

PROPOSITION 2.20. *For any two fractional ideals $\mathfrak{A},\mathfrak{B}$ of L we have*

$$|\mathrm{GL}(\mathfrak{A}\oplus\mathfrak{B})^+\backslash\mathbb{P}^1(L)| = h, \tag{2.73}$$

where h is the class number of L.

PROOF. Consider the map:

$$\mathrm{GL}(\mathfrak{A}\oplus\mathfrak{B})^+\backslash\mathbb{P}^1(L) \longrightarrow Cl(L), \tag{2.74}$$

given by

$$\begin{pmatrix}\alpha\\ \beta\end{pmatrix} \mapsto (\mathfrak{A},\mathfrak{B})\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = \alpha\mathfrak{A}+\beta\mathfrak{B}. \tag{2.75}$$

Using $(\mathfrak{A},\mathfrak{B})\mu = (\mathfrak{A},\mathfrak{B})$ for $\mu \in \mathrm{GL}(\mathfrak{A}\oplus\mathfrak{B})^+$, we see that the map is well-defined. It is surjective because every fractional ideal $\mathfrak{C}$ of L is of the form $\mathfrak{C} = \alpha\mathfrak{A}+\beta\mathfrak{B}$, for some $\alpha,\beta \in L$. Indeed, every ideal class contains infinitely many prime ideals. Take α and β such that $\alpha\mathfrak{A}\mathfrak{C}^{-1}$ and $\beta\mathfrak{B}\mathfrak{C}^{-1}$ are distinct prime ideals. Or if one wishes to use less machinery: assume w.l.o.g. that $\mathfrak{C} = \mathcal{O}_L$ and that $\mathfrak{A}$ is integral. Consider $\mathcal{O}_L/\mathfrak{A} = \sum(\mathcal{O}_{L\mathfrak{p}_i}/{\mathfrak{p}_i}^{e_i}\mathcal{O}_{L\mathfrak{p}_i})$, where $\mathfrak{A} = \prod {\mathfrak{p}_i}^{e_i}$ (Chinese Reminder Theorem). We may assume that $\mathfrak{B}$ is integral. Again by the CRT (weak approximation), we may choose $\beta \in L$ such that for every i we have $\beta\mathfrak{B}\mathcal{O}_{L\mathfrak{p}_i} = \mathcal{O}_{L\mathfrak{p}_i}$ and $\beta\mathfrak{B} \subset \mathcal{O}_L$.

We now show that this map is injective. Suppose that in the class group of L we have

$$\alpha\mathfrak{A} + \beta\mathfrak{B} = \alpha'\mathfrak{A} + \beta'\mathfrak{B}. \tag{2.76}$$

Multiplying ${}^t(\alpha,\beta)$ by a suitable $\lambda \in L^\times$, we may assume that (2.76) is an equality of fractional ideals

$$\alpha\mathfrak{A} + \beta\mathfrak{B} = \alpha'\mathfrak{A} + \beta'\mathfrak{B}. \tag{2.77}$$

Let $\mathfrak{C}^{-1} = \alpha\mathfrak{A} + \beta\mathfrak{B}$. Then

$$\mathcal{O}_L = \alpha\mathfrak{A}\mathfrak{C} + \beta\mathfrak{B}\mathfrak{C} = \alpha'\mathfrak{A}\mathfrak{C} + \beta'\mathfrak{B}\mathfrak{C}. \tag{2.78}$$

Note that $\alpha,\alpha' \in (\mathfrak{A}\mathfrak{C})^{-1}$ and $\beta,\beta' \in (\mathfrak{B}\mathfrak{C})^{-1}$. We may then find $A, A' \in \mathfrak{A}\mathfrak{C}$ and $B, B' \in \mathfrak{B}\mathfrak{C}$ such that

$$1 = \alpha A + \beta B = \alpha' A' + \beta' B'. \tag{2.79}$$

Consider the matrices of determinant 1

$$M = \begin{pmatrix} \alpha & -B \\ \beta & A \end{pmatrix}, \quad N = \begin{pmatrix} \alpha' & -B' \\ \beta' & A' \end{pmatrix}. \tag{2.80}$$

Note that $M\,{}^t(1,0) = {}^t(\alpha,\beta)$ and $N\,{}^t(1,0) = {}^t(\alpha',\beta')$. Thus, $NM^{-1}\,{}^t(\alpha,\beta) = {}^t(\alpha',\beta')$. Now,

$$NM^{-1} = \begin{pmatrix} \alpha'A + \beta B' & \alpha' B - \alpha B' \\ \beta' A - \beta A' & \beta' B + \alpha A' \end{pmatrix}. \tag{2.81}$$

This is a matrix in $\mathrm{GL}(\mathfrak{A}\oplus\mathfrak{B})^+$ (in fact, in $\mathrm{SL}(\mathfrak{A}\oplus\mathfrak{B})^+$), e.g., $\alpha' B - \alpha B' \in (\mathfrak{A}\mathfrak{C})^{-1}\mathfrak{B}\mathfrak{C} + (\mathfrak{A}\mathfrak{C})^{-1}\mathfrak{B}\mathfrak{C} = \mathfrak{A}^{-1}\mathfrak{B}$. □

To make $\mathrm{GL}(\mathfrak{A}\oplus\mathfrak{B})^+\backslash\mathbb{P}^1(L)\coprod\mathrm{GL}(\mathfrak{A}\oplus\mathfrak{B})^+\backslash\mathcal{H}^g$ into a topological space one defines a fundamental system of neighborhoods at $(i\infty,\dots,i\infty)$ by setting:

$$U_r := \{z \in \mathcal{H}^g | \mathrm{Im}(z_i) > r\ \forall i\}, \quad r \in \mathbb{R}. \tag{2.82}$$

Acting on those neighborhoods by elements $\mu \in \mathrm{GL}_2(L)$, we get neighborhoods of all $(\alpha:\beta) \in \mathbb{P}^1(L)$, and this extends the topology to $\mathrm{GL}(\mathfrak{A}\oplus\mathfrak{B})^+\backslash\mathbb{P}^1(L)\coprod\mathrm{GL}(\mathfrak{A}\oplus\mathfrak{B})^+\backslash\mathcal{H}^g$. One then show that in fact this topological space has a structure of a normal compact complex variety. This compactification is called the Satake compactification of $\mathrm{GL}(\mathfrak{A}\oplus\mathfrak{B})^+\backslash\mathcal{H}^g$. See [**36**, Chapter II, p.44].

From the point of view of moduli, one may ask whether the cusps have a moduli interpretation. As in the curve of elliptic curves, the answer is yes. In the case of elliptic curves the moduli interpretation is given using generalized elliptic curves. The precise formulation requires some care, especially when introducing level structure. We refer to [**22**], [**27**].

Ignoring the case of level structure, we may very roughly say that the compactified moduli space classifies semi-abelian varieties with RM by $\mathcal{O}_L$. By a semi-abelian variety $\mathcal{G}$ over a base scheme S we understand a group scheme over S, $\pi : \mathcal{G} \longrightarrow S$, such that π is smooth and every geometric fiber of π is an extension of an abelian variety by a torus. We say that the semi-abelian scheme $\mathcal{G} \longrightarrow S$ has real multiplication if there is given an embedding of rings $\mathcal{O}_L \hookrightarrow \mathrm{End}_S(\mathcal{G})$, such that the Lie algebra of $\mathcal{G} \longrightarrow S$, i.e., the relative tangent space $\mathfrak{t}_{\mathcal{G}/S}$ is a locally free sheaf of $\mathcal{O}_L \otimes_{\mathbb{Z}} \mathcal{O}_S$-modules of rank 1.[3]

Now suppose that one of the geometric fibres of $\mathcal{G} \longrightarrow S$, say $\mathcal{G}_s$ is not an abelian variety. Then we have

$$0 \longrightarrow T \longrightarrow \mathcal{G}_s \longrightarrow A \longrightarrow 0, \tag{2.83}$$

where T is a non-trivial torus and A is an abelian variety. The ring $\mathcal{O}_L$ acts non-trivially on T and therefore L acts non trivially on the rational vector space $X(T) \otimes \mathbb{Q}$, where $X(T)$ are the characters of T. It follows that $\dim(X(T) \otimes \mathbb{Q}) = g$ and therefore $\mathcal{G}_s = T$. That is,

Lemma 2.21. *A semi-abelian variety with RM has fibers that are either tori or abelian varieties.*

We further note, that if the torus T is split then its RM structure is completely determined by the lattice of characters $X(T)$ as an $\mathcal{O}_L$-module. This lattice has rank g and is therefore a projective $\mathcal{O}_L$-module of rank 1. Thus, we expect that there would be h cusps to the compactified moduli space, corresponding to the split tori of dimension g with RM by $\mathcal{O}_L$. This is indeed the case.

Finally we note the following. The stabilizer of $(i\infty, \ldots, i\infty)$ in $\mathrm{SL}(\mathfrak{A} \oplus \mathfrak{B})$ [4] is precisely the matrices $\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$ where a and d are in $\mathcal{O}_L$ and satisfy $ad = 1$, and b is any element the ideal $\mathfrak{D} = \mathfrak{A}^{-1}\mathfrak{B}$. This group is isomorphic to

$${\mathcal{O}_L}^{\times} \ltimes \mathfrak{D}. \tag{2.84}$$

Dividing first only by the action of $\mathfrak{D}$, we get local coordinates as follows: we choose a basis $\nu_1, \ldots, \nu_g$ to the dual of $\mathfrak{D}$ (with respect to the trace pairing). Then

$$x_1 = \exp^{2\pi i \cdot Tr(\nu_1 \tau)}, \ldots, x_g = \exp^{2\pi i \cdot \mathrm{Tr}(\nu_g \tau)}, \tag{2.85}$$

($\tau = (\tau_1, \ldots, \tau_g)$ and $\mathrm{Tr}(\nu\tau) = \sigma_1(\nu)\tau_1 + \cdots + \sigma_g(\nu)\tau_g$) are local coordinates for $\mathfrak{D}\backslash\mathcal{H}^g$ around $(i\infty, \ldots, i\infty)$. Every function f invariant under the action of $\mathfrak{D}$, namely under translation by the elements of $\mathfrak{D}$ (as every modular form w.r.t. $\mathrm{GL}(\mathfrak{A} \oplus \mathfrak{B})^+$ is) has a Taylor expansion

$$f = \sum_{n \in \mathbb{Z}^g} a_n x^n \tag{2.86}$$

(using multi-index notation). But this is the same as writing

$$f = \sum_{\nu \in \mathfrak{D}^\vee} a_\nu \exp^{2\pi i \cdot \mathrm{Tr}(\nu \cdot \tau)}. \tag{2.87}$$

[3] One needs to modify the definition if d_L is not invertible on S. See Chapter 5

[4] The groups $\mathrm{SL}(\mathfrak{A} \oplus \mathfrak{B})$ and $\mathrm{GL}(\mathfrak{A} \oplus \mathfrak{B})^+$ have the same number of cusps and only minor modifications are needed to discuss the case of $\mathrm{GL}(\mathfrak{A} \oplus \mathfrak{B})^+$. The difference eventually amounts to some extra relations satisfied by the coefficients of the q-expansion.

This is the q-expansion we shall define in Chapter 5, Section 2.

3. Hilbert Modular Forms

We keep the notation of the previous sections. Thus L is totally real of degree g, with different $\mathcal{D}_L$, discriminant d_L and ring of integers $\mathcal{O}_L$; $\mathrm{Emb}(L,\mathbb{C}) = \{\sigma_1,\dots,\sigma_g\}$, $h^+ = |Cl(L)^+|$, and $h = |Cl(L)|$. In what follows, $\mathfrak{A}$ and $\mathfrak{B}$ are fractional ideals, and

$$\mathrm{GL}(\mathcal{O}_L \oplus \mathfrak{A})^+ = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a,d \in \mathcal{O}_L, b \in \mathfrak{A}, c \in \mathfrak{A}^{-1}, ad-bc \in {\mathcal{O}_L^\times}^+ \right\}. \tag{3.1}$$

Given a matrix $\delta = \begin{pmatrix} \delta_1 & \delta_2 \\ \delta_3 & \delta_4 \end{pmatrix} \in \mathrm{GL}_2(\mathbb{R})^+$, put

$$j(\delta, z) = (\delta_3 z + \delta_4)(\det \delta)^{-1/2}. \tag{3.2}$$

For a vector $\underline{k} \in \mathbb{Z}^g$, and given $\mu \in \mathrm{GL}_2(L)^+$, put:

$$j_{\underline{k}}(\mu, z) = \prod_{i=1}^{g} j(\sigma_i(\mu), z_i)^{k_i}. \tag{3.3}$$

For $f : \mathcal{H}^g \longrightarrow \mathbb{C}$, put

$$(f|_{\underline{k}}\mu)(z) = j_{\underline{k}}(\mu, z)^{-1} f(\mu z). \tag{3.4}$$

For a group $\Gamma \subseteq \mathrm{GL}(\mathcal{O}_L \oplus \mathfrak{A})^+$ of finite index we make the following

Definition 3.1. A *Hilbert modular form* of weight $\underline{k}$ and level Γ is a holomorphic function

$$f : \mathcal{H}^g \longrightarrow \mathbb{C}, \tag{3.5}$$

such that

$$f|_{\underline{k}}\mu = f, \quad \forall \mu \in \Gamma. \tag{3.6}$$

I.e.,

$$f\Big(\frac{\sigma_1(a)z_1 + \sigma_1(b)}{\sigma_1(c)z_1 + \sigma_1(d)}, \dots, \frac{\sigma_g(a)z_g + \sigma_g(b)}{\sigma_g(c)z_g + \sigma_g(d)}\Big) = \left(\prod_{i=1}^{g} (\sigma_i(c)z_i + \sigma_i(d))^{k_i} \det(\sigma_i(\mu))^{-k_i/2} \right) f(z_1,\dots,z_g). \tag{3.7}$$

And we also require f to be holomorphic at infinity in the following sense:

Let $M = \left\{ a : \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \in \Gamma \right\}$. (Note that $[\mathrm{GL}(\mathcal{O}_L \oplus \mathfrak{A})^+ : \Gamma]\mathfrak{A}^{-1} \subseteq M$, therefore M is a projective $\mathcal{O}_L$-module of rank 1). Then f is invariant under translation by elements of M,

$$f|_{\underline{k}} \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} = f(z_1 + \sigma_1(b), \dots, z_g + \sigma_g(b)) = f(z_1,\dots,z_g), \tag{3.8}$$

and hence possesses a q-expansion:

$$f(z) = \sum_{\nu \in M^\vee} a_\nu e^{2\pi i \mathrm{Tr}(\nu \cdot z)}, \tag{3.9}$$

with $M^\vee = \left\{\ell \in L : \mathrm{Tr}_{L/\mathbb{Q}}(\ell m) \in \mathbb{Z}, \forall m \in M\right\}$, and using the notation $\mathrm{Tr}(\nu \cdot z) = \sigma_1(\nu)z_1 + \cdots + \sigma_g(\nu)z_g$. (When $\Gamma = \mathrm{GL}(\mathcal{O}_L \oplus \mathfrak{A}^{-1})^+$, we have $M \cong \mathfrak{A}^{-1}$ and $M^\vee \cong \mathfrak{A}\mathcal{D}_{L/\mathbb{Q}}^{-1}$; in the case $L = \mathbb{Q}$, $g = 1$, we retrieve:

$$f(\tau) = \sum_{n\in\mathbb{Z}} a_n e^{2\pi i n\tau}, \tag{3.10}$$

with $a_n = 0$ for $n < 0$).

We require, that if g is a modular form of weight $\underline{k}$ and level Γ, then for every $\mu \in \mathrm{GL}_2(L)^+$, $f := g|_{\underline{k}}\mu$ has a q-expansion:

$$f(z) = \sum_{\nu\in M^\vee} a_\nu e^{2\pi i \mathrm{Tr}(\nu\cdot z)}, \tag{3.11}$$

with $a_\nu = 0$ unless $\nu = 0$ or $\nu \gg 0$. Here the module M is the one associated to the group $\mu\Gamma\mu^{-1}$.

Under the topology defined in Remark 2.19, this means that f possesses a holomorphic Taylor expansion at every cusp. The next theorem tells us that in fact for $g > 1$ the holomorphy requirement we imposed is automatic !

THEOREM 3.2. (Köcher's Principle) *Let $f(z) = \sum_{\nu\in M^\vee} a_\nu e^{2\pi i \mathrm{Tr}\nu\cdot z}$ be a modular form of weight $\underline{k}$ and level Γ. Assume $g > 1$, then*

$$a_\nu \neq 0 \implies \nu \gg 0 \text{ or } \nu = 0. \tag{3.12}$$

PROOF. First, a lemma:

LEMMA 3.3. *Put $r_{\underline{k}}(\epsilon) = \prod_{i=1}^g \sigma_i(\epsilon)^{-k_i/2}$, for $\epsilon \in \mathcal{O}_L^{\times +}$. If $\begin{pmatrix} \epsilon & 0 \\ 0 & 1 \end{pmatrix} \in \Gamma$, then*

$$a_\nu = r_{\underline{k}}(\epsilon) a_{\epsilon\nu} \tag{3.13}$$

PROOF. Let $A(\epsilon) = \begin{pmatrix} \epsilon & 0 \\ 0 & 1 \end{pmatrix}$. On the one hand,

$$\begin{aligned} f(\epsilon z) &= f(\sigma_1(\epsilon)z_1, \ldots, \sigma_g(\epsilon)z_g) \\ &= \sum_{\nu\in M^\vee} a_\nu e^{2\pi i \mathrm{Tr}(\epsilon\nu\cdot z)}. \end{aligned} \tag{3.14}$$

On the other hand,

$$\begin{aligned} f(\epsilon z) &= j_{\underline{k}}(A(\epsilon), z) f(z) \\ &= r_{\underline{k}}(\epsilon) \sum_{\nu\in M^\vee} a_\nu e^{2\pi i \mathrm{Tr}\nu\cdot z}. \end{aligned} \tag{3.15}$$

Equating coefficients of $e^{2\pi i \mathrm{Tr}(\epsilon v\cdot z)}$ we are done. □

Assume now that $a_{\nu_0} \neq 0$ for some $\nu_0 \neq 0$ such that ν_0 is not totally positive: without loss of generality, $\sigma_1(\nu_0) < 0$. Find $\epsilon \gg 0$ such that

$$\sigma_1(\epsilon) > 1, \quad \sigma_i(\epsilon) < 1 \quad \text{for all } 2 \leq i \leq g \tag{3.16}$$

and

The existence of such ϵ follows from the fact that Γ is of finite index in $\mathrm{GL}(\mathcal{O}\oplus\mathfrak{A})^+$ and the lattice structure of the units modulo roots of unity.

Consider the terms of $\sum_{\nu\in M^\vee} a_\nu e^{2\pi i \mathrm{Tr}\nu\cdot z}$, parameterized by $\nu_0\epsilon^m, m = 1, 2, \ldots$. Put $z = i$. Then

$$a_{\nu_0\epsilon^m} e^{-2\pi \mathrm{Tr}(\nu_0\epsilon^m)} = a_{\nu_0} r_{\underline{k}}(\epsilon)^{-m} e^{-2\pi \mathrm{Tr}(\nu_0\epsilon^m)}. \tag{3.17}$$

But when $m \longrightarrow \infty$, $e^{-2\pi \mathrm{Tr}(\nu_0\epsilon^m)} \sim e^{-2\pi\sigma_1(\nu_0)\cdot\sigma_1(\epsilon)^m}$, and the exponential growth insures $r_{\underline{k}}(\epsilon)^{-m} e^{-2\pi \mathrm{Tr}(\nu_0\epsilon^m)} \longrightarrow \infty$. So the general term does not approach zero and thus the sum is divergent; contradiction. □

There is a geometrical explanation to the Köcher principle: If Γ is torsion free, a modular form f of weight $\underline{k}$ is a section a of line bundle on $\Gamma\backslash\mathcal{H}^g$. In fact, $j_{\underline{k}}(\mu, z)$ is the factor of automorphy for this line bundle. This is very similar to the case of elliptic curves discussed in Chapter 1, Section 4. One can interprete this line bundle in terms of moduli. Recall that for every $z \in \mathcal{H}^g$ we constructed an abelian variety

$$A_z = \mathbb{C}^g/(\mathcal{O}_L \cdot z + \mathfrak{A}\cdot 1) \tag{3.18}$$

with $\mathcal{O}_L$-action and polarization module. The trivial line bundle

$$\mathbb{C}^g \times \mathcal{H}^g \longrightarrow \mathcal{H}^g \tag{3.19}$$

could be interpreted canonically as either the tangent space $\mathfrak{t}_{A_z,0}$ or the cotangent space $\mathfrak{t}^*_{A_z,0}$ to A_z at the origin. The usual decomposition $\mathbb{C}^g = \oplus_{i=1}^g \mathbb{C}$ is in fact a decomposition of $\mathfrak{t}^*_{A_z,0}$ into 1-dimensional vector spaces $\oplus_{i=1}^g L_{i,z}$ such that $\mathcal{O}_L$ acts on $L_{i,z}$ via σ_i.

Let $\underline{k} = (0, \ldots, 1, \ldots, 0)$ (1 in the i-th place) then $j_{\underline{k}}(\mu, z)$ is the automorphy factor for the line bundle $L_{i,z}$. That is, the relative cotangent space to the "universal" polarized abelian scheme with $(\mathcal{D}_L\mathfrak{A})^{-1}$-polarization $\mathcal{X} \longrightarrow \Gamma\backslash\mathcal{H}^g$, denoted $\mathfrak{t}^*_{\mathcal{X},\Gamma\backslash\mathcal{H}^g}$ decomposes into line bundles: $\mathfrak{t}^*_{\mathcal{X},\Gamma\backslash\mathcal{H}^g} = \oplus_{i=1}^g L_i$ such that $\mathcal{O}_L$ acts on L_i via σ_i.

The transformation law (3.6) shows that a modular form f of weight $\underline{k}$ is a global section of the line bundle determined by the factor of automorphy $j_{\underline{k}}(\mu, z)$. That is, $f \in \Gamma(\Gamma\backslash\mathcal{H}^g, \otimes_{i=1}^g L_i^{\otimes k_i})$.

As discussed above, one may add finitely many cusps[5] to $\Gamma\backslash\mathcal{H}^g$ and get a compact normal complex manifold $\Gamma\backslash\mathcal{H}^{g*}$. Moreover, the line bundle $\otimes_{i=1}^g L_i^{\otimes k_i}$ extends to $\Gamma\backslash\mathcal{H}^{g*}$ and f extends to a meromorphic section of a this line bundle. If f is *not* holomorphic then normality implies that its divisor of poles is of codimension 1 (and thus "visible" on $\Gamma\backslash\mathcal{H}^g$). Thus f holomorphic on $\Gamma\backslash\mathcal{H}^g$ implies that f is holomorphic on $\Gamma\backslash\mathcal{H}^{g*}$ if $g > 1$.

Definition 3.4. A Hilbert modular form f with respect to Γ is a *cusp form* if the constant a_0 in the Fourier expansion of $f|_{\underline{k}}\mu$ is zero for all $\mu \in \mathrm{GL}_2(L)^+$.

Put

$$\mathcal{M}(\mathbb{C}, k, \Gamma) = \text{ complex vector space of modular forms of weight } \underline{\mathrm{k}}, \text{ level } \Gamma. \tag{3.20}$$

and

$$\mathcal{S}(\mathbb{C}, k, \Gamma) = \text{ complex vector space of cusp forms of weight } \underline{\mathrm{k}}, \text{ level } \Gamma. \tag{3.21}$$

Let $\mathcal{M}(\mathbb{C}, \Gamma) = \bigoplus_{\underline{k}\in\mathbb{Z}^g} \mathcal{M}(\mathbb{C}, k, \Gamma)$ and $\mathcal{S}(\mathbb{C}, \Gamma) = \bigoplus_{\underline{k}\in\mathbb{Z}^g} \mathcal{S}(\mathbb{C}, k, \Gamma)$. The ring $\mathcal{M}(\mathbb{C}, \Gamma)$ is a $\mathbb{Z}^g$-graded ring and $\mathcal{S}(\mathbb{C}, \Gamma)$ is a graded ideal of it.

[5] In fact, h points if $\Gamma = \mathrm{GL}(\mathcal{O}_L \oplus \mathfrak{A})^+$ or $\mathrm{SL}(\mathcal{O}_L \oplus \mathfrak{A})^+$.

PROPOSITION 3.5. 1. *If some* $k_i \leq 0$ *and* $k \neq (0, \ldots, 0)$ *then* $\mathcal{M}(\mathbb{C}, k, \Gamma) = \{0\}$.
2. $\mathcal{M}(\mathbb{C}, 0, \Gamma) = \mathbb{C}, \mathcal{M}(\mathbb{C}, 0, \Gamma) = \{0\}$.
3. *If* $\mathcal{S}(\mathbb{C}, k, \Gamma) \neq \mathcal{M}(\mathbb{C}, k, \Gamma)$, *then* $k_1 = \cdots = k_g$.

PROOF. The case $g = 1$ is classical and we shall assume it known, though the argument below could be used to prove this case as well.

We note that we may replace Γ by a smaller group and hence we may assume that Γ has no elliptic elements. We shall use the fact, mentioned above, that for $g > 1$ the nonsingular complex variety $\Gamma\backslash\mathcal{H}^g$ has a compactification $\Gamma\backslash\mathcal{H}^{g*}$ and it is a normal variety to which every modular form extends and thus the divisor of every modular form on $\Gamma\backslash\mathcal{H}^{g*}$ is already visible in $\Gamma\backslash\mathcal{H}^g$.

Part 2 of the Proposition follows easily, because modular forms of weight 0 are identified with holomorphic functions on a normal compact variety and are therefore constant.

Let $f \in \mathcal{M}(\mathbb{C}, k, \Gamma)$, $f \neq 0$. A *twisted diagonal curve* mod Γ' is a curve of the form $\{\mu(z, \ldots, z) : z \in \mathcal{H}\}$ mod Γ' (for μ a fixed element of $\mathrm{GL}_2(L)^+$). Note that letting μ range over $\mathrm{GL}_2(L)^+$, $\{\mu(z, \ldots, z) : z \in \mathcal{H}\}$ is dense in $\mathcal{H}^g$ (Note that $\mathrm{GL}_2(L)^+$ is dense in $\prod_{i=1}^g \mathrm{GL}_2(\mathbb{R})^+$). Thus $\exists \mu \in \mathrm{GL}_2(L)^+$ such that

$$f|_{\underline{k}}\mu \text{ is not zero on } \{(z, \ldots, z) \in \mathcal{H}\} \mod \Gamma'. \tag{3.22}$$

Take $\Gamma' \subseteq \Gamma$ with Γ' commensurable to Γ, such that $f|_{\underline{k}}\mu$ is modular with respect to Γ'. Consider the diagonal map

$$\mathcal{H}/\Gamma'' \xrightarrow{\Phi} \mathcal{H}^g/\Gamma', \tag{3.23}$$

well-defined for some congruence subgroup $\Gamma'' \subset \mathrm{SL}_2(\mathbb{Z})$. Then $\Phi^*(f|_{\underline{k}}\mu)$ is a non-zero modular form of weight $\sum_{i=1}^g k_i$ with respect to Γ''. Hence

$$\sum_{i=1}^{g} k_i \geq 0. \tag{3.24}$$

Consider first the case $\sum_{i=1}^g k_i > 0$ and some k_i, say k_1, is 0. Let $f \in M_{\underline{k}}(\Gamma)$. We will show ultimately that $f = 0$. We have $a_0 = r_{\underline{k}}(\epsilon)a_0$, for $\begin{pmatrix} \epsilon & 0 \\ 0 & 1 \end{pmatrix} \in \Gamma$. Since $\underline{k} \neq \underline{0}$:

$$r_{\underline{k}}(\epsilon) = \prod_{i=2}^{n} (\sigma_i(\epsilon))^{k_i/2} < 1, \tag{3.25}$$

for a suitable ϵ, and that implies that $a_0 = 0$. The same argument holds for $(f|_{\underline{k}}\mu)$ for any $\mu \in \mathrm{GL}_2(L)^+$; that is, under $\mathrm{GL}_2(L)^+$, all the cusps are equivalents, so we just need to consider the q-expansion at one cusp. Thus, f is a cusp form.

Before proving that $f = 0$ we note that the same argument proves Part 3: If $\underline{k}$ is not parallel, then we can always find ϵ such that:

$$r_{\underline{k}}(\epsilon) = \prod_{i=1}^{n} (\sigma_i(\epsilon))^{k_i/2} \neq 1, \tag{3.26}$$

and $\begin{pmatrix} \epsilon & 0 \\ 0 & 1 \end{pmatrix} \in \Gamma$. Hence the constant coefficient of the expansion is zero, and we have a cusp form. That is, if $f \in \mathcal{M}(\mathbb{C}, k, \Gamma)\backslash\mathcal{S}(\mathbb{C}, k, \Gamma)$ then $k_1 = \cdots = k_g$.

Let us finish the proof of Part 1. Recall the formula:

$$\mathrm{Im}(\delta\tau) = \frac{\det(\delta)\mathrm{Im}(\tau)}{|\delta_3\tau+\delta_4|^2}, \quad \delta = \begin{pmatrix} \delta_1 & \delta_2 \\ \delta_3 & \delta_4 \end{pmatrix} \in \mathrm{GL}_2(\mathbb{R})^+. \tag{3.27}$$

Put

$$g(z) = \left(\prod_{i=1}^{g} (\mathrm{Im} z_i)^{k_i/2}\right) f(z). \tag{3.28}$$

- $g(z)$ is *holomorphic* in z_1, $(k_1 = 0)$!
- $|g|$ is Γ-invariant:

(3.29)

$$\begin{aligned} |g(\mu z)| &= \left(\prod_{i=1}^{g} (\mathrm{Im}(\sigma_i(\mu) z_i))^{k_i/2}\right) |f(\mu z)| \\ &= \prod_{i=1}^{g} \left(\frac{\det(\sigma_i(\mu)) \cdot \mathrm{Im}(z_i)}{|\sigma_i(c) z_i + \sigma_i(d)|^2}\right)^{k_i/2} \prod_{i=1}^{g} |\sigma_i(c) z_i + \sigma_i(d)|^{k_i} (\det \sigma_i(\mu))^{-k_i/2} |f(z)| \\ &= |g(z)| \end{aligned}$$

- $|g| = 0$ at the cusps $(a_0 = 0)$.

The two last statements imply that if $|g|$ is bounded on $\mathcal{H}^g$ (because $\Gamma\backslash\mathcal{H}^{g*}$ is compact), hence, by Liouville's theorem, g is constant as a function of z_1. Fix some $(z_1, \ldots, z_g)$. The assumption yields:

$$\forall \mu \in \Gamma \quad |g(z_1, \sigma_2(\mu) z_2, \ldots, \sigma_g(\mu) z_g)| = |g(z)|. \tag{3.30}$$

However, $\{z_1, \sigma_2(\mu) z_2, \ldots, \sigma_g(\mu) z_g)\}_{\mu \in \Gamma}$ is *dense* in $\{z_1\} \times \mathcal{H}^{g-1}$, hence $|g|$ constant implies $|g| = 0$, so $f = 0$.

Consider now the case $\sum_{i=1}^{g} k_i \geq 0$ and some k_i is negative; say $k_1 < 0$. There exists a non-zero modular form G_4 of weight $\underline{4}$ on Γ (see Section 5.1). Consider $\widetilde{f} = f^4 G_4^{-k_1}$. It is a cusp form of weight $\underline{\widetilde{k}} = (\widetilde{k_1}, \ldots, \widetilde{k_g})$, and $\sum \widetilde{k_i} > 0$, $\widetilde{k_1} = 0$. Therefore $\widetilde{f} = 0$, hence $f = 0$. □

Definition 3.6. A Hilbert modular form of level $\Gamma_0(\mathfrak{B})$ and weight $\underline{k}$ is an h_+ tuple

$$(f_1, \ldots, f_{h^+}), \tag{3.31}$$

where each f_i is a modular form of weight $\underline{k}$ with respect to the subgroup $\Gamma(\mathfrak{A}, \mathfrak{B})$ of $\mathrm{GL}_2(\mathcal{O}_L \oplus \mathfrak{A}^{-1})^+$, where

$$\Gamma(\mathfrak{A}, \mathfrak{B}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a, d \in \mathcal{O}_L, b \in \mathfrak{A}^{-1}, c \in \mathfrak{A}\mathfrak{B}, ad - bc \in \mathcal{O}_L^{\times +} \right\}. \tag{3.32}$$

4. More on the diagonal curve

Consider the case $[L:\mathbb{Q}]=2$. To ease notation we assume that $L\subset\mathbb{C}$ and we write $\mathrm{Emb}(L,\mathbb{C})=\{1,\sigma\}$. We have a commutative diagram:

$$\begin{array}{ccc}\mathcal{H} & \longrightarrow & \mathcal{H}^2 \\ \downarrow & & \downarrow \\ \mathcal{H}/\mathrm{SL}_2(\mathbb{Z}) & \xrightarrow{\Phi} & \mathcal{H}^2/\mathrm{SL}_2(\mathcal{O}_L)\end{array}\quad , \tag{4.1}$$

where Φ is the diagonal map: $\Phi(z)=(z,z)$, and $\mathrm{SL}_2(\mathcal{O}_L)=\mathrm{SL}_2(\mathcal{O}_L\oplus\mathcal{O}_L)$.

Proposition 4.1. *The map Φ is generically injective.*

Proof. Let $\lambda\in\mathrm{SL}_2(\mathcal{O}_L)$ be a matrix such that for all $z\in\mathcal{H}$ there exists $\tau\in\mathcal{H}$ such that $\lambda(z,z)=(\tau,\tau)$. Then $\forall z$ we have $\sigma(\lambda)^{-1}\lambda z=z$. So $\sigma(\lambda)^{-1}\lambda=\begin{pmatrix}a&b\\c&d\end{pmatrix}$ and $cz^2+dz=az+b$ for every $z\in\mathcal{H}$. Hence $b=c=0$ and $a=d$, with $ad=1$. That is, $\sigma(\lambda)^{-1}\lambda=\pm1$. We have to consider two cases:

1. $\sigma(\lambda)=\lambda\implies\lambda\in\mathrm{SL}_2(\mathbb{Z})$.
2. $\sigma(\lambda)=-\lambda\implies\lambda=\sqrt{d}\begin{pmatrix}a&b\\c&d\end{pmatrix},\ a,b,c,d\in\mathbb{Z}\implies\det\lambda\neq1$,

whence a contradiction. Thus, the stabilizer of the diagonal in $\mathrm{SL}_2(\mathcal{O}_L)$ is $\mathrm{SL}_2(\mathbb{Z})$ and the proof is complete. □

4.1. Modular interpretation. Let $\tau\in\mathcal{H}$ and let Λ be the lattice $\mathbb{Z}+\mathbb{Z}\tau$ in $\mathbb{C}$; let $\widetilde{\Lambda}$ be the lattice $\mathcal{O}_L\cdot(\tau,\tau)+\mathcal{O}_L\cdot(1,1)$ in $\mathbb{C}^2$. Consider the map

$$\mathbb{C}\longrightarrow\mathbb{C}\otimes\mathcal{O}_L,\quad z\mapsto z\otimes1. \tag{4.2}$$

It induces a map

$$\mathbb{C}/\Lambda\xrightarrow{\Phi}\mathbb{C}\otimes\mathcal{O}_L/\Lambda\otimes\mathcal{O}_L. \tag{4.3}$$

Under the isomorphism $\mathbb{C}\otimes\mathcal{O}_L\cong\mathbb{C}^2$, given by $z\otimes\ell\mapsto{}^t(\ell z,\sigma(\ell)z)$, the lattice $\Lambda\otimes\mathcal{O}_L$ is mapped to the lattice $\widetilde{\Lambda}$. Let $E=\mathbb{C}/\Lambda$. We note that the map Φ is non other then the map

$$E\longrightarrow E\otimes_{\mathbb{Z}}\mathcal{O}_L, \tag{4.4}$$

given by $r\mapsto r\otimes1$.

Assume that $\mathcal{O}_L=\mathbb{Z}\left[\sqrt{d}\right]$, with basis $1,\sqrt{d}$. Then

$$\widetilde{\Lambda}=\left\langle\begin{pmatrix}\tau\\\tau\end{pmatrix},\begin{pmatrix}\sqrt{d}\tau\\-\sqrt{d}\tau\end{pmatrix},\begin{pmatrix}1\\1\end{pmatrix},\begin{pmatrix}\sqrt{d}\\-\sqrt{d}\end{pmatrix}\right\rangle, \tag{4.5}$$

and $\sqrt{d}$ acts on $\mathbb{C}^2$ by $\begin{pmatrix}\sqrt{d}&0\\0&-\sqrt{d}\end{pmatrix}$. The rational representation of $\sqrt{d}$ with respect to the basis given in (4.5) is: $\sqrt{d}\mapsto\begin{pmatrix}0&d&0&0\\1&0&0&0\\0&0&0&d\\0&0&1&0\end{pmatrix}$.

We may change the coordinate on $\mathbb{C}^2$ by:

$$\begin{pmatrix} 1 & \sqrt{d} \\ 1 & -\sqrt{d} \end{pmatrix}^{-1} = -\frac{1}{2\sqrt{d}} \begin{pmatrix} -\sqrt{d} & -\sqrt{d} \\ -1 & 1 \end{pmatrix}. \tag{4.6}$$

Under that change of coordinates $\widetilde{\Lambda}$ is mapped to

$$\widetilde{\widetilde{\Lambda}} = \left\langle \begin{pmatrix} \tau \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \tau \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle, \tag{4.7}$$

(This lattice is obviously defining $E \times E$) and $\sqrt{d}$ acts now on $\mathbb{C}^2$ by

$$\begin{pmatrix} 1 & \sqrt{d} \\ 1 & -\sqrt{d} \end{pmatrix}^{-1} \begin{pmatrix} \sqrt{d} & 0 \\ 0 & -\sqrt{d} \end{pmatrix} \begin{pmatrix} 1 & \sqrt{d} \\ 1 & -\sqrt{d} \end{pmatrix} = \begin{pmatrix} 0 & d \\ 1 & 0 \end{pmatrix} \tag{4.8}$$

(with the same rational representation).

What are the $\mathcal{O}_L$-linear symmetric homomorphisms $\mu : E \times E \longrightarrow (E \times E)^\vee$?

Let $\lambda : E \times E \longrightarrow (E \times E)^\vee$ be the canonical polarization. That is, if we let $p_i : E \times E \longrightarrow E$ be the projection on the i-th component then $\lambda = \phi_{\mathcal{L}}$, where $\mathcal{L} = p_1^* \mathcal{O}_E([0]) \otimes p_2^* \mathcal{O}_E([0])$.

Remark 4.2. The map μ is $\mathcal{O}_L$-linear iff $\mu m = m^\vee \mu$ for all $m \in \mathcal{O}_L$, iff $(\lambda^{-1}\mu)m = (\lambda^{-1} m^\vee \lambda)(\lambda^{-1}\mu)$.

Recall that $\mu \longmapsto \lambda^{-1}\mu$ gives an embedding $\mathrm{NS}^0 \hookrightarrow \mathrm{End}^0(A)^{sym}$, that $m \longmapsto \lambda^{-1} m^\vee \lambda$ is the Rosati involution defined by λ. In our case, identifying $\mathrm{End}(E^2)$ with $M_2(\mathrm{End}(E))$, where $\mathrm{End}(E) = \mathbb{Z}$ or an order in an imaginary quadratic field, one finds that the Rosati involution is given by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \longmapsto \begin{pmatrix} a^\vee & c^\vee \\ b^\vee & d^\vee \end{pmatrix}, \tag{4.9}$$

where $^\vee$ is complex conjugation. Thus, the symmetric elements are the elements of the form

$$\begin{pmatrix} \alpha & \beta \\ \beta^\vee & \delta \end{pmatrix}, \tag{4.10}$$

with $\alpha, \delta \in \mathbb{Z}$ and $\beta \in \mathrm{End}(E)$.

Write $\lambda^{-1}\mu = \begin{pmatrix} \alpha & \beta \\ \beta^\vee & \delta \end{pmatrix}$, take $m = \begin{pmatrix} 0 & d \\ 1 & 0 \end{pmatrix}$. The map μ is $\mathcal{O}_L$-linear iff

$$\begin{pmatrix} \alpha & \beta \\ \beta^\vee & \delta \end{pmatrix} \begin{pmatrix} 0 & d \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ d & 0 \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ \beta^\vee & \delta \end{pmatrix}. \tag{4.11}$$

That is, if and only if

$$\begin{pmatrix} \beta & d\alpha \\ \delta & d\beta^\vee \end{pmatrix} = \begin{pmatrix} \beta^\vee & \delta \\ d\alpha & d\beta \end{pmatrix}. \tag{4.12}$$

So $\beta = \beta^\vee$, $\delta = \alpha d$, and this implies:

$$\mathcal{M}_A = \left\{ \begin{pmatrix} \alpha & \beta \\ \beta & d\alpha \end{pmatrix} : \alpha, \beta \in \mathbb{Z} \right\}, \tag{4.13}$$

$$\mathcal{M}_A^+ = \left\{ \begin{pmatrix} \alpha & \beta \\ \beta & d\alpha \end{pmatrix} \in \mathcal{M}_A : \alpha > 0, d\alpha^2 - \beta^2 > 0 \right\}. \tag{4.14}$$

But $\mathcal{D}_L^{-1} = \left\langle \frac{1}{2}, \frac{1}{2\sqrt{d}} \right\rangle$ and $\mathcal{M}_A \cong \mathcal{D}_L^{-1}$ as $\mathcal{O}_L$-modules via

$$\begin{pmatrix} \alpha & \beta \\ \beta & d\alpha \end{pmatrix} \longmapsto \alpha \cdot \frac{1}{2} + \beta \cdot \frac{1}{2\sqrt{d}}, \tag{4.15}$$

because

$$\begin{pmatrix} \alpha & \beta \\ \beta & d\alpha \end{pmatrix} \begin{pmatrix} 0 & d \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \beta & d\alpha \\ d\alpha & d\beta \end{pmatrix} \longmapsto \tag{4.16}$$

$$\beta \cdot \frac{1}{2} + d\alpha \cdot \frac{1}{2\sqrt{d}} = (\alpha \cdot \frac{1}{2} + \beta \cdot \frac{1}{2\sqrt{d}})\sqrt{d}.$$

Under this isomorphism,

$$\begin{aligned} \mathcal{M}_A^+ &\longmapsto \left\{ \alpha \cdot \frac{1}{2} + \beta \frac{1}{2\sqrt{d}} : \alpha > 0, d\alpha^2 > \beta^2 \right\} \\ &= \left\{ \alpha \cdot \frac{1}{2} + \beta \frac{1}{2\sqrt{d}} : \alpha + \frac{\beta}{\sqrt{d}} > 0, \alpha - \frac{\beta}{\sqrt{d}} > 0 \right\} \\ &= (\mathcal{D}_L^{-1})^+. \end{aligned} \tag{4.17}$$

Note that this fits with $\mathrm{SL}_2(\mathcal{O}_L)\backslash\mathcal{H}^2 = \mathrm{SL}_2(\mathcal{O}_L \oplus \mathcal{O}_L)\backslash\mathcal{H}^2$ parameterizing abelian varieties, constructed from lattices $\mathcal{O}_L \cdot \tau + \mathcal{O}_L$, and having polarization module $(\mathcal{D}_L^{-1}, (\mathcal{D}_L^{-1})^+)$.

5. Construction Of Hilbert Modular Forms

5.1. Eisenstein series. Fix a fractional ideal $\mathfrak{A}$ of L.

The complex manifold

$$\mathrm{GL}(\mathcal{O}_L \oplus \mathfrak{A}^{-1})^+ \backslash \mathcal{H}^g. \tag{5.1}$$

parameterizes complex abelian varieties with RM by $\mathcal{O}_L$ and polarization module $(\mathfrak{A}\mathcal{D}_L^{-1}, (\mathfrak{A}\mathcal{D}_L^{-1})^+)$. Take $\mathfrak{B}$ a fractional ideal of L with class B in $Cl(L)$. Put $\underline{k} = (k, \ldots, k)$, $k \geq 2, k \in 2\mathbb{Z}$.

Definition 5.1. The Eisenstein series of weight $\underline{k}$ and class B is:

$$G_{\underline{k},B}(z) := \mathbb{N}(\mathfrak{B})^k \sum_{(\alpha,\beta) \in \mathfrak{B}\mathfrak{A} \oplus \mathfrak{B}}'' \mathbb{N}(\alpha z + \beta)^{-k}. \tag{5.2}$$

where the symbol $''$ indicates the following: We say that (α, β) and (μ, δ) are *associated* if there exists an $\epsilon \in \mathcal{O}_L^\times$ such that:

$$(\alpha, \beta) = (\epsilon\mu, \epsilon\delta). \tag{5.3}$$

This is an equivalence relation, and the summation is over representatives of equivalence classes different from $\{(0,0)\}$. We use the notation $\mathbb{N}(v) = \mathbb{N}(v_1, \ldots, v_g) = v_1 \cdots v_g$.

Let us show that an Eisenstein series of weight $\underline{k}$ and class B in indeed well-defined: If $(\alpha, \beta) \sim (\mu, \delta)$, say $(\alpha, \beta) = (\epsilon\mu, \epsilon\delta)$, the norms satisfy

$$\mathbb{N}(\alpha z + \beta)^{-k} = \mathbb{N}(\epsilon)^{-k}\mathbb{N}(\mu z + \delta)^{-k} = \mathbb{N}(\mu z + \delta)^{-k}, \tag{5.4}$$

because $\underline{k}$ is parallel and k is even implies $\mathbb{N}(\epsilon)^{-k} = 1$.

REMARK 5.2. The sum converges for $k > 2$, and also for $k = 2$, provided $g > 1$, and defines a holomorphic function on $\mathcal{H}^g$. Note that $G_{\underline{k},B}$ depends only on the ideal class of B.

An Eisenstein series is modular form with respect to $\mathrm{GL}(\mathcal{O}_L \oplus \mathfrak{A}^{-1})^+$. To see that note first that the group $\mathrm{GL}_2(\mathcal{O}_L \oplus \mathfrak{A}^{-1})^+$ is contained in the automorphism group of $\mathfrak{B}\mathfrak{A} \oplus \mathfrak{B}$:

$$(\mathfrak{B}\mathfrak{A} \oplus \mathfrak{B})\mu = \mathfrak{B}\mathfrak{A} \oplus \mathfrak{B}, \ \ \forall \mu \in \mathrm{GL}_2(\mathcal{O}_L \oplus \mathfrak{A}^{-1})^+. \tag{5.5}$$

This follows from the definition of $\mathrm{GL}_2(\mathcal{O}_L \oplus \mathfrak{A}^{-1})^+$ as invertible elements with totally positive determinant of the ring

$$\begin{pmatrix} \mathcal{O}_L & \mathfrak{A}^{-1} \\ \mathfrak{A} & \mathcal{O}_L \end{pmatrix}. \tag{5.6}$$

The action of $\mathrm{GL}_2(\mathcal{O}_L \oplus \mathfrak{A}^{-1})^+$ respect the equivalence relation we defined on pairs $(\alpha, \beta) \in \mathfrak{B}\mathfrak{A} \oplus \mathfrak{B}$. Furthermore, let $\mu = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ then

$$\begin{aligned} \mathbb{N}(\alpha\mu z + \beta)^k &= [\mathbb{N}(\alpha(az+b) + \beta(cz+d))][\mathbb{N}(cz+d)^{-1}]^{-k} \\ &= \mathbb{N}\left((\alpha,\beta)\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} z \\ 1 \end{pmatrix}\right)^{-k} \mathbb{N}(cz+d)^k. \end{aligned} \tag{5.7}$$

Thus

$$G_{k,B}(\mu z) = \mathbb{N}((cz+d))^k G_{k,B}(z) = j_k(\mu, z) G_{k,B}(z), \tag{5.8}$$

where, as usual,

$$j_k(\mu, z) = \mathbb{N}((cz+d))^k \mathbb{N}(\det \mu)^{-k/2} = \mathbb{N}((cz+d))^k. \tag{5.9}$$

DEFINITION 5.3. Let A be an ideal class, $\mathfrak{b}$ an integral ideal. Put

$$\zeta_A(k) = \sum_{\mathfrak{c} \in A, \mathfrak{c} \subseteq \mathcal{O}_L} \mathbb{N}(\mathfrak{c})^{-k}; \qquad \sigma_{k-1,A}(\mathfrak{b}) = \sum_{\mathfrak{c} \in A, \mathfrak{b} \subseteq \mathfrak{c} \subseteq \mathcal{O}_L} \mathbb{N}(\mathfrak{c})^{k-1}. \tag{5.10}$$

THEOREM 5.4. 1. *The modular forms $G_{k,B}$ on $\mathrm{GL}(\mathcal{O}_L \oplus \mathfrak{A}^{-1})^+ \backslash \mathcal{H}^g$ are linearly independent over $\mathbb{C}$, and span a vector space of dimension $h = |Cl(L)|$.*

2. *The Fourier expansion of $G_{k,B}$ with respect to the cusp $(i\infty, \ldots, i\infty)$ is*

$$G_{k,B} = c\left\{ \frac{\zeta_{\mathcal{D}_L B}(1-k)}{2^g} + \sum_{\nu \in \mathfrak{A}\mathcal{D}_L^{-1}, \nu \gg 0} \sigma_{k-1, B\mathcal{D}_L}((\nu)\mathfrak{A}^{-1}\mathcal{D}_L) e^{2\pi i \mathrm{Tr}\nu z} \right\}, \tag{5.11}$$

with $c = \frac{(2\pi i)^{kg}}{(k-1)!^g} d_L^{1/2-k}$.

REMARK 5.5. The number $\frac{\zeta_{\mathcal{D}_L B}(1-k)}{2^g}$ is a rational number by a theorem of Siegel, and $\sigma_{k-1, B\mathcal{D}_L}((\nu)\mathfrak{A}^{-1}\mathcal{D}_L)$ are integers by definition.

In what follows, we give a heuristic explanation for the rationality of $\zeta_{\mathcal{D}_L B}(1-k)$:

The algebraic variety V whose underlying analytic variety (formed of its complex points) is isomorphic analytically to $\mathrm{GL}(\mathcal{O}_L \oplus \mathfrak{A}^{-1})^+ \backslash \mathcal{H}^g$ is defined over $\mathbb{Q}$ and so is the space of modular forms of parallel weight $\underline{k}$. In fact if $\mathcal{X} \longrightarrow V$ is the universal object then the space of modular forms is $\Gamma(V, \det \mathfrak{t}_{\mathcal{X},V}^{*\otimes k})$. This vector

space is contained in $\mathrm{Hom}(V, \det \mathfrak{t}^*_{\mathcal{X},V})$, a space with a natural $G(\mathbb{C}/\mathbb{Q})$-action since both $\mathfrak{A}$ and $\det \mathfrak{t}^*_{\mathcal{X},V}$ are defined over $\mathbb{Q}$.

If f is a modular form, so is f^σ for $\sigma \in \mathrm{Gal}(\mathbb{C}/\mathbb{Q})$. The q-expansion principle says that the q-expansion of f is in some sense algebraic. It is really the development of f to power series with respect to local parameters in the local ring at infinity that are defined over the rationals. That is, if $f = \sum_{\nu \in \mathfrak{A}\mathcal{O}_L^{-1}} a_\nu e^{2\pi i \mathrm{Tr}\nu z}$, we expect that $f^\sigma = \sum_{\nu \in \mathfrak{A}\mathcal{O}_L^{-1}} a_\nu^\sigma e^{2\pi i \mathrm{Tr}\nu z}$. Thus, the constant

$$\left(\frac{\zeta_{B,\mathcal{D}_L}(1-k)}{2^g}\right) - \left(\frac{\zeta_{B,\mathcal{D}_L}(1-k)}{2^g}\right)^\sigma = \left(\frac{1}{c}G_{k,B}\right) - \left(\frac{1}{c}G_{k,B}\right)^\sigma \tag{5.12}$$

is a modular form of weight $k > 0$, which is possible only if this constant is zero.

Exercise 5.6. Prove that a modular form of weight $k \neq (0, \dots, 0)$ which is not zero has infinitely many terms in its q-expansion.

We also define the following Eisenstein series:

$$E_k^{L,*} = \frac{1}{c} \sum_{B \in Cl(L)} G_{k,B}, \tag{5.13}$$

and

$$E_k^L = 2^g \zeta_L(1-k)^{-1} E_k^{L,*}. \tag{5.14}$$

Thus, the q-expansion of $E_k^{L,*}$ at the cusp $(i\infty, \dots, i\infty)$ is

$$\frac{\zeta_L(1-k)}{2^g} + \sum_{\nu \in \mathfrak{A}\mathcal{D}_L^{-1}, \nu \gg 0} \sigma_{k-1}((\nu)\mathfrak{A}^{-1}\mathcal{D}_L) e^{2\pi i \mathrm{Tr}\nu z}, \tag{5.15}$$

where

$$\sigma_{k-1}(\mathfrak{b}) = \sum_{\mathfrak{b} \subseteq \mathfrak{c} \subseteq \mathcal{O}_L} \mathbb{N}(\mathfrak{c})^{k-1}. \tag{5.16}$$

5.2. Other methods of constructing modular forms. We indicate more methods of constructing modular forms. In essence they all come one way or another from theta series. The first method, that we now demonstrate by a particular example, is to pull back modular forms from other modular varieties.

Let $\mathfrak{A}$ be a fractional ideal of L, taken with its natural notion of positivity, and consider the (coarsely) representable functor Φ of abelian varieties with RM and polarization module $\mathfrak{A}$. That is, the functor assigning to a scheme S the isomorphism classes of triples over S

$$(A, \iota, m : (\mathcal{M}_A, \mathcal{M}_A^+) \cong (\mathfrak{A}, \mathfrak{A}^+)). \tag{5.17}$$

Choose an element $\lambda \in \mathfrak{A}^+$, consider it as an element of $\mathcal{M}_A^+$, and let $\underline{d} = (d_1 | \dots | d_g)$ be the elementary divisors of $\mathrm{Ker}(\lambda)$. There is a natural transformation of functors from the functor Φ to the (coarsely) representable functor Ψ of abelian varieties of dimension g with a polarization having elementary divisors $\underline{d} = (d_1, \dots, d_g)$. If we let $\mathcal{M}^{\mathfrak{A}}(\mathbb{C})$ denote the moduli scheme over $\mathbb{C}$ that (coarsely) represents Φ, and we let $\mathcal{S}_{\underline{d}}(\mathbb{C})$ denote the moduli scheme that (coarsely) represents Ψ, [6] we get a

[6] Scheme like $\mathcal{M}^{\mathfrak{A}}$ usually go under the name of *Hilbert modular schemes* (not to be confused with Hilbert schemes!), while the schemes $\mathcal{S}_{\underline{d}}$ usually go under the name of Siegel modular schemes.

morphism of schemes

$$f : \mathcal{M}^{\mathfrak{A}}(\mathbb{C}) \longrightarrow \mathcal{S}_{\underline{d}}(\mathbb{C}). \tag{5.18}$$

(Such morphisms and their generalizations, are usually called modular embeddings, though they are rarely *embeddings*. Another example of such a map is the diagonal curve discussed in Section 4 and Section 4.1). One may write (5.18) explicitely. See [**106**, pp. 640-1]. It can be formulated as coming from a suitable embedding of $\mathrm{GL}_2(L)$ in $\mathrm{Sp}(2g, \mathbb{Q})$, and the interpretation of the moduli spaces as double cosets spaces of the adelic points of these algebraic groups.

Every modular form of weight k on $\mathcal{S}_{\underline{d}}(\mathbb{C})$ (that means: $f : \mathcal{H}_g \longrightarrow \mathbb{C}$ such that

$$f(\gamma\tau) = \det(C\tau + D)^k f(\tau), \;\; \gamma = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \tag{5.19}$$

in a suitable subgroup of $\mathrm{Sp}(2g, \mathbb{Z})$) pulls back to a modular form on $\mathcal{M}^{\mathfrak{A}}(\mathbb{C})$ of parallel weight $(k, \ldots, k)$. While this may be checked by laborious calculation, using the explicit form of the map $f : \mathcal{M}^{\mathfrak{A}}(\mathbb{C}) \longrightarrow \mathcal{S}_{\underline{d}}(\mathbb{C})$, it may also be justified easily using Katz view of modular forms. We discussed that view point in the case of elliptic curves in Chapter 1, Section 4. The case of abelian varieties with RM is similar and is discussed in Chapter 5, Section **??**.

Based on that we argue as follows. Given

$$(A, \iota : \mathcal{O}_L \hookrightarrow \mathrm{End}_S(A), m : (\mathcal{M}_A, \mathcal{M}_A^+) \longrightarrow (\mathfrak{A}, \mathfrak{A}^+), \omega)/S \tag{5.20}$$

where ω is an $\mathcal{O}_L \otimes \mathcal{O}_S$ basis to $\mathfrak{t}^*_{A/S}$, we get the data (A, λ, ω'), with ω' a basis for $\det \mathfrak{t}^*_{A/S}$. Here, we first decompose $\mathcal{O}_L \otimes \mathcal{O}_S$ as $\oplus_{\sigma \in \mathrm{Emb}(L,\mathbb{C})} \mathcal{O}_S$ ($\mathcal{O}_S$ is a sheaf of $\mathbb{C}$-algebras), and get from ω an $\mathcal{O}_S$ basis, $\{\omega_1, \ldots, \omega_g\}$ to $\mathfrak{t}^*_{A/S}$. We then let $\omega' = \omega_1 \wedge \cdots \wedge \omega_g$. Thus, the rule is

$$(A, \iota, m, \omega)/S \longmapsto f(A, \lambda, \omega'). \tag{5.21}$$

It is immediate that this is a modular form!

Take for example the case when $\mathfrak{A} = \mathcal{O}_L$ and λ is the principal polarization corresponding to 1. Then we consider a map

$$\mathrm{GL}(\mathcal{O}_L \oplus \mathcal{D}^{-1}_{L/\mathbb{Q}})^+ \backslash \mathcal{H}^g \longrightarrow \mathrm{Sp}(2g, \mathbb{Z}) \backslash \mathcal{H}_g. \tag{5.22}$$

On $\mathcal{H}_g$ one can define Riemann's theta functions with characteristic $\left[\begin{smallmatrix} \epsilon \\ \epsilon' \end{smallmatrix} \right] \in \mathbb{Q}^{2g}$ by

$$\Theta\left[\begin{smallmatrix} \epsilon \\ \epsilon' \end{smallmatrix} \right](\tau) = \sum_{N \in \mathbb{Z}^g} \exp\left(2\pi i \left\{ \frac{1}{2} \, {}^t(N + \epsilon)\tau(N + \epsilon) + \, {}^t(N + \epsilon)\epsilon' \right\} \right). \tag{5.23}$$

They are holomorphic functions on $\mathcal{H}_g$. One can show, using the transformation formula for theta functions ([**67**, Chapter 8.6]) that if $n\epsilon, n\epsilon' \in \mathbb{Z}^g$ then $\Theta^{8n}\left[\begin{smallmatrix} \epsilon \\ \epsilon' \end{smallmatrix} \right]$ is a modular form of weight $4n$ on the principal congruence subgroup of level $2n$; if n is even $\Theta^{2n}\left[\begin{smallmatrix} \epsilon \\ \epsilon' \end{smallmatrix} \right]$ is a modular form of weight n on the principal congruence subgroup of level n; $\Theta^2\left[\begin{smallmatrix} \epsilon \\ \epsilon' \end{smallmatrix} \right]$ is a modular form of weight 1 on the principal congruence subgroup of level $4n^2$.

By pulling back these modular forms we get Hilbert modular forms, of some level. We note that the construction is arithmetic. The morphism

$$f : \mathcal{M}^{\mathcal{O}_L} \longrightarrow \mathcal{S}_{(1,\dots,1)} \tag{5.24}$$

exists over any scheme. The theta functions often have the g.c.d. of their coefficients equal to one, and in any event the g.c.d. divides $2n$ if $n\epsilon, n\epsilon' \in \mathbb{Z}$.

We remark that Shimura used vector valued modular forms in [**106**, pp. 640-642] to obtain Hilbert modular forms of weight $(1, \dots, 1, 2, 1, \dots, 1)$.

EXERCISE$^\bigstar$ 5.7. For which weights are there non-zero modular forms? (Same question in positive characteristic).

6. Siegel's formula

Consider the case $\mathfrak{A} = \mathcal{O}_L$. Let

$$\Phi : \mathrm{SL}_2(\mathbb{Z})\backslash\mathcal{H} \longrightarrow \mathrm{SL}_2(\mathcal{O}_L)\backslash\mathcal{H} \tag{6.1}$$

be the diagonal curve. Let $F_k = \Phi^* E_k^{L,*}$. It is a modular form on $\mathrm{SL}_2(\mathbb{Z})$ of weight gk with q-expansion:

$$F_k = \frac{\zeta_L(1-k)}{2^g} + \sum_{n=1}^{\infty} a_k(n) q^n, \tag{6.2}$$

where

$$a_k(n) = \sum_{\substack{\nu \in (\mathcal{D}_L^{-1})^+ \\ \mathrm{Tr}(\nu) = n}} \sigma_{k-1}((\nu)\mathcal{D}_L) = \sum_{\substack{\nu \in (\mathcal{D}_L^{-1})^+ \\ \mathrm{Tr}(\nu) = n}} \sum_{(\nu)\mathcal{D}_L \subseteq \mathfrak{A} \subseteq \mathcal{O}_L} \mathbb{N}(\mathfrak{A})^{k-1}. \tag{6.3}$$

Lets apply this when $g = k = 2$. Since the space of modular forms on $SL_2(\mathbb{Z})$ of weight 4 is one dimensional, F_2 is a multiple of

$$G_{4,\mathrm{SL}_2(\mathbb{Z})} = \frac{\zeta_{\mathbb{Q}}(-3)}{2} + \sum_{n=1}^{\infty} c_3(n) q^n = \frac{1}{240} + q + \cdots. \tag{6.4}$$

Hence $\frac{1}{4}\zeta_L(-1)/\frac{1}{240} = a_2(1)$, and

$$\zeta_L(-1) = \frac{1}{60} \sum_{\substack{\nu \in (\mathcal{D}_L^{-1})^+ \\ \mathrm{Tr}(\nu) = 1}} \sum_{(\nu)\mathcal{D}_L \subseteq \mathfrak{A} \subseteq \mathcal{O}_L} \mathbb{N}(\mathfrak{A}). \tag{6.5}$$

LEMMA 6.1. *We have the following identity:*

$$a_2(1) = \sum_{\substack{a \in \mathbb{Z},\, |a| < \sqrt{d_L} \\ a \equiv d_L \pmod 2}} \sigma_1\left(\frac{d_L - a^2}{4}\right). \tag{6.6}$$

PROOF. We have

$$\mathcal{D}_L^{-1} = \left\{ \frac{1}{\sqrt{d_L}} \left(\frac{a + b\sqrt{d_L}}{2} \right) : a, b \in \mathbb{Z}, a, b \equiv d_L \pmod 2 \right\} = \frac{1}{\sqrt{d_L}} \mathcal{O}_L. \tag{6.7}$$

EXERCISE 6.2. Write the totally positive elements ν of trace 1 in $\mathcal{D}_L^{-1}$ explicitly. Find out the ideals $\nu\mathcal{D}_L$ and obtain the formula

$$a_2(1) = \sum_{\substack{|a| < \sqrt{d_L} \\ a \equiv d_L \pmod 2}} \sum_{(\frac{a+\sqrt{d_L}}{2}) \subseteq \mathfrak{A} \subseteq \mathcal{O}_L} \mathbb{N}(\mathfrak{A}). \tag{6.8}$$

We rewrite (6.8) as

$$a_2(1) = \sum_{\substack{|a| < \sqrt{d_L} \\ a \equiv d_L \pmod 2}} \sum_{d \mid \frac{d_L - a^2}{4}} d \cdot S_{a,d_L}(d), \tag{6.9}$$

where

$$S_{a,d_L}(d) = \sum_{(\frac{a+\sqrt{d_L}}{2}) \subseteq \mathfrak{A} \subseteq \mathcal{O}_L,\, \mathbb{N}\mathfrak{A} = d} 1. \tag{6.10}$$

We claim that $S_{a,d_L}(d) = 1$. Write

$$(\frac{a+\sqrt{d_L}}{2}) = \prod \mathcal{S}_i^{a_i} \overline{\mathcal{S}}_i^{c_i} \prod \mathcal{R}_i^{\beta_i} \prod \mathcal{T}_i^{\mu_i}, \tag{6.11}$$

where the first product run over split primes, the second over ramified primes and the last one over inert primes.

Recall that the primes dividing $\sqrt{d_L}$ are precisely the ramified primes. If $\mathrm{val}_{\mathcal{T}_i}(\frac{a+\sqrt{d_L}}{2}) > 0$ then $\mathrm{val}_{\mathcal{T}_i}(\frac{a-\sqrt{d_L}}{2}) > 0$, which implies that $\mathrm{val}_{\mathcal{T}_i}(\sqrt{d_L})$, being equal to $\mathrm{val}_{\mathcal{T}_i}(\frac{a+\sqrt{d_L}}{2} + \frac{a-\sqrt{d_L}}{2})$ is positive. Contradiction. If both $\mathrm{val}_{\mathcal{S}_i}(\frac{a-\sqrt{d_L}}{2}) = \mathrm{val}_{\overline{\mathcal{S}_i}}(\frac{a+\sqrt{d_L}}{2})$, and $\mathrm{val}_{\mathcal{S}_i}(\frac{a+\sqrt{d_L}}{2})$ are positive, then $\mathrm{val}_{\mathcal{S}_i}(\sqrt{d_L}) > 0$. Again, a contradiction. We conclude that we may write:

$$\left(\frac{a+\sqrt{d_L}}{2}\right) = \prod \mathcal{S}_i^{a_i} \prod \mathcal{R}_i^{\beta_i}, \quad \text{for } i \neq j, \mathcal{S}_i \neq S_j, \overline{S_j}. \tag{6.12}$$

Therefore

$$\frac{d_L - a^2}{4} = \prod s_i^{a_i} \prod r_i^{\beta_i}, \quad (s_i) = \mathcal{S}_i \overline{\mathcal{S}_i}, (r_i) = \mathcal{R}_i^2. \tag{6.13}$$

The inclusion $\frac{a+\sqrt{d_L}}{2} \subseteq \mathfrak{A}$, $\mathbb{N}\mathfrak{A} = d$ implies,

$$d = \prod s_i^{a_i'} \prod r_i^{\beta_i'}, \quad a_i' \leq a_i, \beta_i' \leq \beta_i. \tag{6.14}$$

And therefore

$$\mathfrak{A} = \prod \mathcal{S}_i^{a_i'} \prod \mathcal{R}_i^{\beta_i'}, \tag{6.15}$$

That is, $\mathfrak{A}$ is *uniquely* determined by d and the condition $(\frac{a+\sqrt{d_L}}{2}) \subseteq \mathfrak{A}$, $\mathbb{N}\mathfrak{A} = d$, hence the claim.

□

COROLLARY 6.3. **Siegel's formula**

$$\zeta_L(-1) = \frac{1}{60} \sum_{\substack{a \in \mathbb{Z},\, a \equiv d_L \pmod 2 \\ |a| < \sqrt{d_L}}} \sigma_1\left(\frac{d_L - a^2}{4}\right). \tag{6.16}$$

EXERCISE 6.4. (See [**13**]) Prove

$$\zeta_K(-3) = \frac{1}{120} \sum_{\substack{a \in \mathbb{Z},\, a \equiv d_L \pmod 2 \\ |a| < \sqrt{d_L}}} \sigma_3\left(\frac{d_L - a^2}{4}\right). \tag{6.17}$$

CHAPTER 3

Abelian Varieties with Real Multiplication over General Fields

In this chapter we explain some methods of studying abelian varieties in positive characteristic. After defining abelian varieties over a general field, the dual abelian variety, polarization and the Weil pairing, we turn to study finite Heisenberg groups. These groups, introduced by Mumford, allow one to study line bundles on abelian varieties and their behaviour under isogenies and are replacing, in the situation of characteristic p, the more powerful tool given by the Appell-Humbert theorem.

We discuss various methods to understand abelian varieties over a field of positive characteristic: The Honda-Tate method that describes isogeny classes of abelian varieties over finite fields; Serre-Tate coordinates that describe ordinary abelian varieties and their deformations over an algebraically closed field of finite characteristic, and Deligne's refinement. Following Chai-Norman and Ribet, we give some applications to moduli spaces.

1. Abelian Varieties over a General Field

In this section, we will develop some general features of abelian varieties defined over any field k. Recall that an *abelian variety* A over k is a connected, reduced projective algebraic group.

LEMMA 1.1. (Rigidity lemma) *Let $f : V \times W \longrightarrow U$ be a morphism of algebraic varieties over k. Assume*

1. *V is projective.*
2. *There exist $v_0 \in V, w_0 \in W$ such that*

$$f(\{v_0\} \times W) = f(V \times \{w_0\}) = \{u_0\}\,. \tag{1.1}$$

Then $f(V \times W) = \{u_0\}$.

PROOF. Let $U_0 \subseteq U$ be an open affine set such that $u_0 \in U_0$, let $\pi_W : V \times W \longrightarrow W$ be the projection and let

$$Z = \pi_W(f^{-1}(U \backslash U_0)). \tag{1.2}$$

The set $U \backslash U_0$ is closed. Since V is projective, hence proper over k, π_W is a closed map, and thus Z is closed. The set Z is also characterized by by the following property:

$$w \notin Z \iff f(V \times \{w\}) \subseteq U_0. \tag{1.3}$$

Hence $W \backslash Z$ is a non empty, hence dense, open set in W. For $w \notin Z$, $f(V \times \{w_0\})$ is a point, being the the image of a projective variety in an affine subset. Combining this with

$$f(\{v_0\} \times W) = \{u_0\}\,, \tag{1.4}$$

we get,

$$f(V \times (W \backslash Z)) = \{u_0\}\,. \tag{1.5}$$

Since $V \times (W\backslash Z)$ is dense in $V \times W$, we conclude: $f(V \times W) = \{u_0\}$. □

COROLLARY 1.2. *Every morphism $f : A \longrightarrow B$ of abelian varieties over k is of the form $T_b \circ h$ where $h : A \longrightarrow B$ is a homomorphism of abelian varieties, $b \in B$, and T_b is the translation-by-b-map: $T_b(b') = b' + b$.*

PROOF. Replacing f by $T_b \circ f$, with $b = -f(0)$, we may assume $f(0) = 0$ and we need to prove that f is a homomorphism. We apply the rigidity lemma to the morphism

$$F : A \times A \longrightarrow B, \tag{1.6}$$

given by

$$F = f \circ m_A - m_B \circ (f, f). \tag{1.7}$$

(I.e., on points, $F(x, y) = f(x + y) - f(x) - f(y)$).

The assumptions of the rigidity lemma hold: $F(A \times \{0\}) = 0_B = F(\{0\} \times A)$. Hence F is the constant map with image 0_B. I.e., f is a homomorphism. □

COROLLARY 1.3. *The group law on an abelian variety is commutative.*

PROOF. Apply Corollary 1.2 to the inverse morphism:

$$inv : A \longrightarrow A. \tag{1.8}$$

It gives that *inv* is a homomorphism, hence the group law is commutative. □

COROLLARY 1.4. *There is a unique group law on A for which 0_A is the identity.*

PROOF. Apply the corollary 1.2 to the identity map ! □

One can prove that the multiplication by n map, denoted $[n]$ (or simply n), is an isogeny of degree n^{2g}. That is, it is a surjective homomorphism and $A[n]$, the Kernel of $[n]$, defined by the cartesian diagram:

$$\begin{array}{ccc} A[n] & \xrightarrow[\text{immersion}]{\text{closed}} & A \\ \downarrow & & \downarrow{\scriptstyle [n]} \\ \mathrm{Spec}(k) & \longrightarrow & A \end{array}, \tag{1.9}$$

is an affine group scheme of order n^{2g} (where $g = \dim A$), i.e. $A[n] = \mathrm{Spec}(R)$ where R is a Hopf algebra of dimension n^{2g}. If char(k) = 0, or $(\mathrm{char(k)}, n) = 1$, then $A[n]$ is étale, i.e., $A[n] \otimes_k k^s \cong (\mathbb{Z}/n\mathbb{Z})^{2g}$. If char(k) = p and $n = p$, then

$$|A[p](\overline{k})| \leq p^g. \tag{1.10}$$

By a theorem of Oort [**94**, (15.11)] every finite commutative group scheme over a perfect field embeds as a subgroup of some abelian variety.

EXERCISE 1.5. Write the automorphism group as a semi direct product of translations and group automorphisms. Find the connected component of this algebraic group. Deduce that it needs not be of finite type.

Use the notion of polarization to find subgroups of finite order. What Galois representations can you obtain this way?

1.1. The dual abelian variety. Let $\mathrm{Pic}(A)$ denote the isomorphism classes of line bundles on $A \otimes \overline{k}$. It is an abelian group under the operation of tensor product. For every line bundle $\mathcal{L}$ the map:

$$\Phi_{\mathcal{L}} : A \longrightarrow \mathrm{Pic}(A), \tag{1.11}$$

$$\Phi_{\mathcal{L}}(x) = T_x^*\mathcal{L} \otimes \mathcal{L}^{-1}, \tag{1.12}$$

is a homomorphism of groups:

THEOREM 1.6. (Theorem of the Square) *We have:*

$$T_{x+y}^*\mathcal{L} \otimes \mathcal{L}^{-1} \cong T_x^*\mathcal{L} \otimes \mathcal{L}^{-1} \otimes T_y^*\mathcal{L} \otimes \mathcal{L}^{-1}. \tag{1.13}$$

PROOF. See [**72**, Theorem 4.5] □

Let $\mathrm{Pic}^0(A)$ denote the isomorphism classes of line bundles $\mathcal{L}$ such that $\Phi_{\mathcal{L}} \equiv 0$. Note that for all $\mathcal{L}$, $\mathrm{Im}(\Phi_{\mathcal{L}}) \subseteq \mathrm{Pic}^0(A)$:

$$T_y^*(T_x^*\mathcal{L} \otimes \mathcal{L}^{-1}) \otimes (T_x^*\mathcal{L} \otimes \mathcal{L}^{-1})^{-1} = T_{x+y}^*\mathcal{L} \otimes T_y^*\mathcal{L}^{-1} \otimes T_x^*\mathcal{L}^{-1} \otimes \mathcal{L} = 0. \tag{1.14}$$

Note that $\mathrm{Pic}^0(A)$ is a subgroup of $\mathrm{Pic}(A)$. This follows from the formula: $\Phi_{\mathcal{L}\otimes\mathcal{M}} = \Phi_{\mathcal{L}} + \Phi_{\mathcal{M}}$. One defines the Néron-Severi group , $\mathrm{NS}(A)$, by

$$\mathrm{NS}(A) = \mathrm{Pic}(A)/\mathrm{Pic}^0(A). \tag{1.15}$$

Then:

$$0 \longrightarrow A^\vee \longrightarrow \mathrm{Pic}(A) \longrightarrow NS(A) \longrightarrow 0. \tag{1.16}$$

EXERCISE 1.7. Let $k = \mathbb{C}$. Use Appell-Humbert theorem to show $\mathrm{Pic}^0(A)$ is precisely the dual abelian variety we defined previously.

One denotes $\mathrm{Pic}^0(A)$ also by $A^\vee$, or A^t, and calls it the *dual abelian variety*. The name is justified by the following theorem and further properties we list below. The proofs of these properties will take us too much astray, and we content ourselves with referring to [**76**], especially Chapters 8 and 13.

THEOREM 1.8. *There exists an abelian variety $A^\vee/k$ such that $A^\vee(\overline{k})$ is in isomorphism (as groups) with* $\mathrm{Pic}^0(A)$.

Moreover, on $A \times A^\vee$ there is a line bundle $\mathcal{P}$, unique up to isomorphism, called the Poincaré bundle*, with the following properties:*

1. *For every $\alpha \in A^\vee$ the line bundle $\mathcal{P}|_{A\times\{\alpha\}}$ represents that element of* $\mathrm{Pic}^0(A)$ *given by α.*
2. *$\mathcal{P}|_{\{0\}\times A^\vee}$ is trivial.*

FACT 1.9.
1. If $\mathcal{L}$ is ample, then $\Phi_{\mathcal{L}} : A \longrightarrow A^\vee$ is an isogeny, i.e. a surjective homomorphism with finite kernel (but the converse is not necessarily true).
2. $A \cong (A^\vee)^\vee$ canonically. The isomorphism is given in fact by $t \mapsto \mathcal{P}|_{t\times A^\vee}$. Moreover, the Poincaré bundle on $A \times A^\vee$, when A is interpreted as $A^{\vee\vee}$, is the same bundle.
3. Let f_1, f_2 be homomorphisms and $f_1^\vee, f_2^\vee : B^\vee \longrightarrow A^\vee$ the corresponding homomorphisms of the dual abelian varieties ($f_i^\vee(\mathcal{L})$ is just the pull-back $f_i^*(\mathcal{L})$). Then it is a theorem that $(f_1 + f_2)^\vee = f_1^\vee + f_2^\vee$.

EXERCISE 1.10. Use the fact that $A^\vee$ is an abelian variety to deduce the theorem of the square from Corollary 1.2.

THEOREM 1.11. *Let*

$$0 \longrightarrow H \longrightarrow A \xrightarrow{f} B \longrightarrow 0 \tag{1.17}$$

be an exact sequence with A, B abelian varieties, and H a finite group scheme. Then

$$0 \longrightarrow H^\vee \longrightarrow B^\vee \xrightarrow{f^\vee} A^\vee \longrightarrow 0 \tag{1.18}$$

where $f^\vee(\mathcal{L}) = f^*(\mathcal{L})$, and $H^\vee$ is the group scheme dual to H, i.e., representing the functor $S \mapsto \mathrm{Hom}(H(S), \mathbb{G}_m(S))$.

We refer to Appendix A for properties of group schemes.

PROOF. See [**75**, Theorem 1, Section 15, Chapter III, p. 143] □

REMARK 1.12. The duality theory of abelian varieties and the special role of the Poincaré bundle can be developed much further. In particular, there is a Fourier transform. See [**84**], [**25**].

REMARK 1.13. The relation $f \circ f^\vee = [\deg(f)]$ holds for elliptic curves, but does not hold in general if $\dim(A) > 1$.

DEFINITION 1.14. We define a *polarization* of A to be a homomorphism,

$$f : A \longrightarrow A^\vee, \tag{1.19}$$

such that $f = \Phi_{\mathcal{L}}$ for some *ample* line bundle $\mathcal{L}$.

FACT 1.15. If f is a polarization, then f is an isogeny and $f = f^\vee$ (under $(A^\vee)^\vee \cong A$). Hence, the kernel of a polarization is a self-dual group scheme.

The polarization is called *principal* if f is an isomorphism. Since A is a variety, $A \subseteq \mathbb{P}^N$, for some N, by definition. Thus A possesses an ample line bundle. In particular, A always carries some polarization and is isogenous to its dual $A^\vee$.

For every commutative group scheme H over k, there exists a perfect pairing

$$H \times H^\vee \longrightarrow \mathbb{G}_m. \tag{1.20}$$

Now,

$$0 \longrightarrow A[n] \longrightarrow A \xrightarrow{n} A \longrightarrow 0,, \tag{1.21}$$

gives:

$$0 \longrightarrow (A[n])^\vee \longrightarrow A^\vee \xrightarrow{n^\vee} A^\vee \longrightarrow 0. \tag{1.22}$$

On the other hand, note that $1^\vee = 1$, and since $(f+g)^\vee = f^\vee + g^\vee$, it follows by induction that $n^\vee = n$. Thus we conclude that

$$A^\vee[n] = A[n]^\vee, \tag{1.23}$$

and we therfore obtain a perfect pairing.:

$$A[n] \times A[n]^\vee \longrightarrow \mathbb{G}_m. \tag{1.24}$$

If $\lambda : A \longrightarrow A^\vee$ is a polarization, we get a bilinear, antisymmetric, $\mathrm{Gal}(\overline{k}/k)$-invariant pairing:

$$\langle\ ,\ \rangle_\lambda : A[n] \times A[n] \longrightarrow \mathbb{G}_m, \tag{1.25}$$

$$\langle x, y\rangle_\lambda = \langle x, \lambda(y)\rangle\,. \tag{1.26}$$

This pairing is called the *Weil pairing*. It is perfect iff $(\deg\lambda, n) = 1$.

We mentioned in Section 1 that if $\mathrm{char}(k) = p > 0$ then

$$|A[p](\overline{k})| \leq p^g. \tag{1.27}$$

One may prove this by showing that if A is isogenous to B, then

$$|A[p](\overline{k})| = |B[p](\overline{k})|. \tag{1.28}$$

Take $B = A^\vee$, and use the Weil pairing:

$$A[p] \times A[p]^\vee \longrightarrow \mathbb{G}_m. \tag{1.29}$$

If we put $H = A[p](\overline{k})$, $K = A^\vee[p](\overline{k})$. The map

$$H \times K \longrightarrow \mathbb{G}_m(\overline{k}) \tag{1.30}$$

factors through μ_p and we get:

$$H \times K \longrightarrow \mu_p(\overline{k}) = 1. \tag{1.31}$$

Hence $K \subseteq H^\perp$. The perfectness implies $|K| \leq p^{2g}/|H|$. On the other hand, $|K| = |H|$, thus $|H| \leq p^g$.

2. Finite Heisenberg Groups

Over the complex numbers, the Appell-Humbert theorem allows one complete understanding of line bundles on abelian varieties and thus of polarizations and their behaviour under isogenies. The Appell-Humbert theorem rests on the fact that we have a surjective analytic map $\mathbb{C}^g \longrightarrow A$ where $\mathbb{C}^g$ is a contractible space that is independent of A. In characteristic p, there is no similar construction known. Thus one needs other methods to manage line bundles and polarizations. This is provided by the finite Heisenberg group defined by Mumford. For an extensive treatment, see [**76**] and [**67**]. See also [**82**].

Let A be an abelian variety over an algebraically closed field k, and $\mathcal{L}$ an *ample* line bundle on A. We assume, for simplicity, that either $\mathrm{char}(k) = 0$, or $(\deg\mathcal{L}, \mathrm{char}(k)) = 1$.

Definition 2.1. Let: $G(\mathcal{L}) := \{\phi | \phi : \mathcal{L} \longrightarrow \mathcal{L}\}$ such that ϕ is an isomorphism of $\mathcal{L}$ that covers a translation map on the base and induces a linear map on fibers. That is, there exists $a \in A$ such that the following diagram commutes and ϕ is linear on fibers:

$$\begin{array}{ccc} \mathcal{L} & \xrightarrow{\phi} & \mathcal{L} \\ \downarrow & & \downarrow \\ A & \xrightarrow{T_a} & A \end{array}\,. \tag{2.1}$$

The group law is composition: $\phi_2 \circ \phi_1$ is given by

$$\begin{array}{ccccc} \mathcal{L} & \xrightarrow{\phi_1} & \mathcal{L} & \xrightarrow{\phi_2} & \mathcal{L} \\ \downarrow & & \downarrow & & \downarrow \\ A & \xrightarrow{T_{a_1}} & A & \xrightarrow{T_{a_2}} & A \end{array}\,. \tag{2.2}$$

One calls $G(\mathcal{L})$ the *Heisenberg group*, or *Theta group*, associated to $(A, \mathcal{L})$.

An equivalent way of defining $G(\mathcal{L})$ is as the set of pairs (ψ, a), where $a \in A$ and $\psi : \mathcal{L} \longrightarrow T_a^*\mathcal{L}$ is an isomorphism. Indeed, let $\rho : \mathcal{L} \longrightarrow A$ be the projection. Then $T_a^*\mathcal{L} = \{(x, l) : \rho(l) = T_a(x) = x + a\}$. The map $\psi : \mathcal{L} \longrightarrow T_a^*\mathcal{L}$ has the form $(\rho(\ell), \ell) \mapsto (\rho(\ell), \phi(\ell))$, for some ϕ, where $\rho(\phi(\ell)) = \rho(\ell) + a$. Thus,

$$\begin{array}{ccc} \mathcal{L} & \xrightarrow{\phi} & \mathcal{L} \\ \downarrow{\scriptstyle \rho} & & \downarrow{\scriptstyle \rho} \\ A & \xrightarrow{T_a} & A \end{array} . \tag{2.3}$$

Conversely, given ϕ such that (2.3) holds, define $\psi(\rho(\ell), \ell) = (\rho(\ell), \phi(\ell))$. We note that in this language we have $(\phi_2, a_2) \circ (\phi_1, a_1) = (T_{a_1}^*\phi_2 \circ \phi_1, a_1 + a_2)$.

The group $G(\mathcal{L})$ sits in an exact sequence:

$$0 \longrightarrow \mathbb{G}_m \longrightarrow G(\mathcal{L}) \xrightarrow{\pi} K(\mathcal{L}) \longrightarrow 0. \tag{2.4}$$

The group $K(\mathcal{L})$ is defined as $\{a : T_a^*\mathcal{L} \cong \mathcal{L}\} = \mathrm{Ker}(\Phi_{\mathcal{L}})$. Fixing a, ϕ is unique up to a scalar. To see that, note that if ϕ' is another such map, then

$$\phi \circ \phi'^{-1} : \mathcal{L} \longrightarrow \mathcal{L} \tag{2.5}$$

is an isomorphism of line bundles, and thus for every $a \in A$ the morphism $f : \mathcal{L}_a \longrightarrow \mathcal{L}_a$ is multiplication by a constant $f(a)$ and $f(a) \neq 0, \infty$. We get therefore a morphism $f : A \longrightarrow \mathbb{G}_m$. But A is projective and $\mathbb{G}_m$ is affine, hence f is constant.

We have the following:

FACT 2.2. $\deg(\mathcal{L}) = \dim_k \Gamma(A, \mathcal{L})$, and $|K(\mathcal{L})| = (\deg(\mathcal{L}))^2$.

One finds immediately that $\mathbb{G}_m \subset Z(G(\mathcal{L}))$, the center of the Heisenberg group, and we may define an alternating bilinear pairing, the *Mumford pairing*

$$e^{\mathcal{L}} : K(\mathcal{L}) \times K(\mathcal{L}) \longrightarrow \mathbb{G}_m, \tag{2.6}$$

by

$$e^{\mathcal{L}}(x, y) = [\widetilde{x}, \widetilde{y}], \tag{2.7}$$

where $\widetilde{x}, \widetilde{y}$ are lifts of x and y, respectively, to $G(\mathcal{L})$, and $[a, b] = aba^{-1}b^{-1}$.

DEFINITION 2.3. A subgroup $H \subseteq K(\mathcal{L})$ is *isotropic* if $e^{\mathcal{L}}(H, H) = 1$.

DEFINITION 2.4. A subgroup $\widetilde{H}$ of $G(\mathcal{L})$ is called a *level subgroup* if $\pi(\widetilde{H}) \cong \widetilde{H}$. If $H = \pi(\widetilde{H})$, we say $\widetilde{H}$ *is above* H .

EXERCISE 2.5. There exists a level subgroup $\widetilde{H}$ above H iff H is isotropic.

Let $f : A \longrightarrow B$ be an isogeny of abelian varieties, $\mathcal{M}$ an ample line bundle on B. Let $\mathcal{L} = f^*\mathcal{M}$. Then we have a canonical level subgroup $H_{\mathcal{M}}$ above $H = \mathrm{Ker}f$. It is constructed as follows: Given $a \in \mathrm{Ker}f$,

$$T_a^*\mathcal{L} \cong T_a^*f^*\mathcal{M} \cong (f \circ T_a)^*\mathcal{M} \cong f^*\mathcal{M} \cong \mathcal{L}. \tag{2.8}$$

All isomorphisms being canonical. This yields a canonical isomorphism $\phi_a : \mathcal{L} \cong T_a^*\mathcal{L}$, for every $a \in H$ and $H_{\mathcal{M}} = \{\phi_a : a \in H\}$ is a level subgroup of $G(\mathcal{L})$ above H. Conversely, given a level subgroup $\widetilde{H} \subset G(\mathcal{L})$, let $H = \pi(\mathcal{L})$ and consider:

$$\begin{array}{ccc} \mathcal{L} & \longrightarrow & \mathcal{L}/\widetilde{H} \\ \downarrow & & \downarrow \\ A & \xrightarrow{p_H} & A/H \end{array} . \tag{2.9}$$

This is an inverse construction:

$$H_{\mathcal{L}/\widetilde{H}} = \widetilde{H}. \tag{2.10}$$

The construction is in fact a manifestation of the theory of descent. Using it one shows that there is bijection between the level subgroups over H and the line bundles $\mathcal{M}$ on A/H such that $p_H^*(\mathcal{M}) = \mathcal{L}$. This is the true content of the two exact sequences:

$$0 \longrightarrow H \longrightarrow A \xrightarrow{p_H} A/H \longrightarrow 0, \tag{2.11}$$

and

$$0 \longrightarrow H^\vee \longrightarrow \mathrm{Pic}(A/H) \xrightarrow{p_H^*} \mathrm{Pic}(A) \longrightarrow 0\cdot \tag{2.12}$$

Note that any two level subgroups H_1, H_2 over H differ by a character χ of H: Let $x \in H$ and x', x'' be its lifts to H_1 and H_2 respectively. Let $\chi(x) = (x')^{-1}(x'')$. Then χ is a character. Conversely, if $\chi \in H^\vee$ and H_1 is a level subgroup over H define H_2 by $\{x \cdot \chi(\pi(x)) : x \in H_1\}$.

LEMMA 2.6. *Let $f : A \longrightarrow B$ be an isogeny with kernel H, $\mathcal{M}$ an ample line bundle on B, and $\mathcal{L} = f^*\mathcal{M}$. Then*

$$K(\mathcal{M}) \cong H^\perp/H. \tag{2.13}$$

PROOF. Let $H_{\mathcal{M}}$ be the level subgroup corresponding to $\mathcal{M}$ defined above. To prove the claim, consider the set of all possible liftings of isomorphisms

$$\begin{array}{ccc} \mathcal{M} & \xrightarrow{\phi} & \mathcal{M} \\ \downarrow & & \downarrow \\ B & \xrightarrow{T_b} & B \end{array} \tag{2.14}$$

to elements of $G(\mathcal{L})$. It is easy to see that for $\phi \in G(\mathcal{M})$ the lifting $p_H^*\phi$ is an element of $G(\mathcal{L})$ commuting with $H_{\mathcal{M}}$, and $p_H^*\phi$ is well-defined up to a choice of an element of $\mathrm{Ker}(p_H) = H$. Indeed, given $a \in A$ such that $f(a) = b$ there is a unique morphism ϕ' completing the following diagram:

$$\begin{array}{ccc} p_H^*\mathcal{M} & \xrightarrow{\phi'} & p_H^*\mathcal{M} \\ \downarrow & & \downarrow \\ A & \xrightarrow{T_a} & A \end{array} \qquad \xrightarrow{f} \qquad \begin{array}{ccc} \mathcal{M} & \xrightarrow{\phi} & \mathcal{M} \\ \downarrow & & \downarrow \\ B & \xrightarrow{T_b} & B \end{array} . \tag{2.15}$$

Conversely, every automorphism in $\pi^{-1}(H^\perp) = \mathrm{Cent}_{G(\mathcal{L})}(H_{\mathcal{M}})$ will descend to an automorphism of $\mathcal{M}$. Thus, in fact $G(\mathcal{M}) \cong \mathrm{Cent}_{G(\mathcal{L})}(H_{\mathcal{M}})/H_{\mathcal{M}}$ and hence $K(\mathcal{M}) \cong H^\perp/H$. □

FACT 2.7.

$$\deg \mathcal{M} = \deg \mathcal{L}/|\mathrm{Ker} f|. \tag{2.16}$$

This also follows from descent theory: $\Gamma(B, \mathcal{M}) \cong \Gamma(A, f^*\mathcal{M})^{H_{\mathcal{M}}}$.

COROLLARY 2.8. 1. *We have the following equalities:*

$$(\deg \mathcal{M})^2 = |K(\mathcal{M})| = |H^\perp|/|H| = (\deg \mathcal{L}/|H|)^2 = |K(\mathcal{L})|/|H|^2. \tag{2.17}$$

2. *We have* $|K(\mathcal{L})| = |H| \cdot |H^\perp|$. *Hence,* $e^{\mathcal{L}}$ *is a perfect pairing and* $\mathbb{G}_m = Z(G(\mathcal{L}))$.

COROLLARY 2.9. *Every abelian variety* A/k *is isogenous to a principally polarized abelian variety.*

PROOF. Let $\mathcal{L}$ be an ample line bundle on A. Let $H \subset K(\mathcal{L})$ be a maximal isotropic subgroup.[1] Let $\widetilde{H}$ be a level subgroup above H. Let $\mathcal{M}$ be the line bundle $\mathcal{L}/\widetilde{H}$ on A/H. Clearly (A, ϕ_L) is isogenous to $(A/H, \phi_{\mathcal{M}})$. Moreover, by Corollary 2.8, $\deg(\mathcal{M}) = |H^\perp/H| = 1$. □

There is a connection between the Mumford pairing $e^{\mathcal{L}}$ and the Weil pairing: Assume that $\mathcal{M}$ is an ample line bundle on A and $\mathcal{M} = \mathcal{L}^{\otimes n}$. Then,

$$\begin{aligned} x \in K(\mathcal{M}) &\iff T_x^*\mathcal{M} \cong \mathcal{M} \\ &\iff (T_x^*\mathcal{L})^n \cong \mathcal{L}^n \\ &\iff (T_x^*\mathcal{L} \otimes \mathcal{L}^{-1})^n \cong 0 \\ &\iff n \cdot \phi_{\mathcal{L}}(x) = 0 \\ &\iff x \in [n]^{-1}K(\mathcal{L}) \end{aligned} \tag{2.18}$$

Therefore, if $\mathcal{L}$ is of degree 1, then $K(\mathcal{L}) = \{0\}$ and $K(M) = A[n]$. One can prove that Mumford's pairing $e^{\mathcal{M}}$ and Weil's pairing $<,>_{\phi_L}$ are equal:

$$e^{\mathcal{M}}(x, y) =< x, y >_{\phi_{\mathcal{L}}}, \quad x, y \in A[n]. \tag{2.19}$$

If one accepts (2.19) then one obtains that the $e^{\mathcal{M}}$ pairing is Galois equivariant and depends only on the class of $\mathcal{M}$ mod Pic^0, and, on the other hand, that the Weil pairing is an alternating pairing. The proof of (2.19) is not hard but requires going in detail into the identification of $H^\vee$ as the kernel of $\mathrm{Pic}(A/H) \longrightarrow \mathrm{Pic}(A)$.

The finite Heisenberg groups are instrumental in the construction of moduli spaces for abelian varieties and in studying isogenies between them. Over the complex numbers this translates to identities between Riemann theta functions. An exhaustive treatment may be found in Mumford's trilogy, [**80**], [**81**], [**82**]. Our discussion below is intended mainly to whet the reader's appetite.

Suppose that $\mathcal{L}$ is a very ample line bundle on A. Then a choice of basis $s_0, \ldots, s_r$ of $\Gamma(A, \mathcal{L})$ gives a projective embedding

$$A \longrightarrow \mathbb{P}^r, \tag{2.20}$$

$$a \mapsto (s_0(a) : \cdots : s_r(a)). \tag{2.21}$$

For a general variety A there is no canonical way to choose the basis. In the case of abelian varieties, one can use the fact that $\Gamma(X, \mathcal{L})$ is an irreducible representation

[1] We only discussed those notions in the case where the degree of $\mathcal{L}$ is prime to the characteristic of k. However, the theory may be extended to cover the general case.

of $G(\mathcal{L})$ to try and narrow down the choices of bases, making such a choice "almost canonical".

DEFINITION 2.10. A representation ρ of $G(\mathcal{L})$ into GL_n is called of *weight* 1 if for every $x \in \mathbb{G}_m, \rho(x)$ acts as multiplication by x.

THEOREM 2.11. (Stone-Von Neumann) *Under the assumptions above, there exists a unique irreducible representation of weight* 1. [2]

PROOF. Let V be such a representation, and choose a maximal isotropic subgroup H of $K(\mathcal{L})$. Let $\widetilde{H}$ be a level subgroup above H. We may decompose V according to characters of $\widetilde{H}$

$$V = \bigoplus_{\chi \in \widetilde{H}^*} V^\chi. \tag{2.22}$$

We denote the summand corresponding to the trivial character by $V^{\mathbf{1}}$. Let

$$\kappa : G(\mathcal{L}) \longrightarrow \widetilde{H}^*, \tag{2.23}$$

be defined by

$$\kappa(g) = [\cdot, g]. \tag{2.24}$$

Let $v \in V^\chi$ and $g \in G(\mathcal{L})$. Then for $h \in \widetilde{H}$,

$$\begin{aligned} h(gV) = (hg)V &= ([h,g]gh)V \\ &= [h,g]\chi(h)gV \\ &= (\kappa(g)\chi)(h) \cdot gV. \end{aligned} \tag{2.25}$$

Hence

$$gV^\chi = V^{\kappa(g)\cdot\chi}. \tag{2.26}$$

Therefore, if one $V^\chi \neq 0$, then every $V^\chi \neq 0$ and they are all isomorphic as vector spaces. Let $v \in V^{\mathbf{1}}, v \neq 0$. Then (2.25) implies that $\mathrm{Span}\{gv : g \in G(\mathcal{L})\}$ is a sub representation W of $G(\mathcal{L})$ whose intersection with $V^{\mathbf{1}}$ is $k \cdot v$. Hence $W = V$ and $V^{\mathbf{1}}$ is one dimensional.

Choose a section $[g_\psi : \psi \in \widetilde{H}^*]$ to κ; and choose a basis v to the one dimensional vector space $V^{\mathbf{1}}$. Then

$$V = \bigoplus_{\psi \in \widetilde{H}^*} g_\psi k \cdot v. \tag{2.27}$$

Every element in $G(\mathcal{L})$ can be written uniquely as xyz, with $x \in k^\times, y \in \{g_\psi\}$ and $z \in \widetilde{H}$. ($\mathrm{Ker}(\kappa) = k^\times \cdot \widetilde{H}$, because H was chosen to be maximal.) We see that the representation is completely determined (2.27). The only thing that matters is $zg_\psi = rg_{\psi'}z$ for some ψ'. We remark that in fact (2.27) shows that $V \cong \mathrm{Ind}_{k^\times \widetilde{H}}^G k$ (where k is the trivial one dimensional representation). □

REMARK 2.12. **Almost canonical bases.**

[2] This theorem was greatly generalized by Weil [**121**], [**122**], who used it to derive and generalize Siegel's theorem on quadratic equations. A theorem now known as the Siegel-Weil theorem.

We now apply the representation theory of $G(\mathcal{L})$ to geometry. First note that $\Gamma(A,\mathcal{L})$ is a representation of $G(\mathcal{L})$. The action being described by the following diagram:

$$\begin{array}{ccc} \mathcal{L} & \xrightarrow{\phi} & \mathcal{L} \\ \uparrow{\scriptstyle s} & & \uparrow{\scriptstyle \phi\cdot s} \\ A & \xleftarrow[T_a^{-1}]{} & A \end{array} \tag{2.28}$$

That is,

$$\phi \cdot s = \phi \circ s \circ T_a^{-1}. \tag{2.29}$$

The center of $G(\mathcal{L})$ acts by scalars and $\dim \Gamma(X,\mathcal{L})$ has the dimension of the unique irreducible representation of weight 1. Therefore $\Gamma(A,\mathcal{L})$ *is* that unique irreducible representation.

Let $\mathcal{L}$ be a symmetric ample line bundle. That is, $[-1]^*\mathcal{L} \cong \mathcal{L}$. This induces a canonical automorphism ι of $G(\mathcal{L})$, where $\iota(\phi) = \iota^*\phi$.

Let $d_1,\dots,d_g$ with $1|d_1|d_2|\cdots|d_g$ be the elementary divisors of $K(\mathcal{L})$, i.e. $K(\mathcal{L}) \cong \oplus(\mathbb{Z}/d_i\mathbb{Z})^2$. One can show $G(\mathcal{L}) \cong G_{(d_1,\dots,d_g)}$, where

$$G_{(d_1,\dots,d_g)} = k^\times \times (\oplus_{i=1}^g \mathbb{Z}/d_i\mathbb{Z}) \times (\oplus_{i=1}^g \mathbb{Z}/d_i\mathbb{Z})^*, \tag{2.30}$$

and

$$(\alpha,x,\phi)(\beta,y,\psi) = (\alpha\beta\psi(x), x+y, \phi\psi). \tag{2.31}$$

An isomorphism $\theta : G(\mathcal{L}) \longrightarrow G_{(d_1,\dots,d_g)}$ restricting to the identity on $\mathbb{C}^\times$, and satisfying $\theta\circ\iota = [-1]\circ\theta$ (where $[-1]$ on the right hand side is the automorphism $G_{(d_1,\dots,d_g)}$ given by $(\alpha,x,\phi)\mapsto(\alpha,-x,\phi^{-1})$) is called a *theta level structure.* . If $\mathcal{M}$ is an ample symmetric line bundle of degree one and $\mathcal{L} = \mathcal{M}^{\otimes n}$, a theta level structure on $G(\mathcal{L})$ is intermediate between full level n structure and full symplectic level $2n$ structure (see [**76**]).

One may consider the moduli of triples $(A,\mathcal{L},\theta)$. Here A is abelian variety (or, more generally, an abelian scheme), $\mathcal{L}$ is an ample line bundle of fixed elementary divisors $d_1,\dots,d_g$, and $\theta : G(\mathcal{L}) \longrightarrow G_{(d_1,\dots,d_g)}$ is a theta level structure. Because of the relation to the usual level structures, proving the representability of such a functor is virtually equivalent to proving the representability of the usual moduli functors.

Given θ, there is an evident choice of level subgroup $\widetilde{H}$, the one corresponding to $\{(1,x,0)\}$, and there is an evident choice of a section to κ, the one corresponding to $\{(1,0,\psi)\}$. Then choosing an non-zero element (unique up to scalar) Θ of $\Gamma(A,\mathcal{L})^{\widetilde{H}}$, we get a basis $\{g_\psi\Theta\}$ to $\Gamma(A,\mathcal{L})$, where the g_ψ are representatives to the cosets of $k^\times\widetilde{H}$ in $G(\mathcal{L})$ corresponding to $\{(1,0,\psi)\}$. The bases obtained by this method are "almost canonical"; they depend on the finitely many choices for θ.

Define a map

$$A \longrightarrow \mathbb{P}(\Gamma(A,\mathcal{L})), \tag{2.32}$$

by

$$a \longmapsto ((g_\psi\Theta)(a))_{\psi\in H^*}. \tag{2.33}$$

According to Mumford, when $\mathcal{L}$ is ample enough (e.g., an 8-th power of an ample symmetric line bundle of degree 1), this embeds our moduli space as a quasi-projective variety. The point corresponding to $(A, \mathcal{M}, \theta)$ is $((g_\psi\Theta)(0_A))$, where 0_A is the identity. More than that, this construction actually gives the universal family over the moduli space.

Over $\mathbb{C}$, for the line bundle considered in the discussion of the Appell-Humbert theorem, those $g_\psi\Theta$ are, up to a common exponent factor, Riemann's theta functions $\Theta\left[\begin{smallmatrix}\epsilon\\ \epsilon'\end{smallmatrix}\right]$:

$$\Theta\left[\begin{smallmatrix}\epsilon\\ \epsilon'\end{smallmatrix}\right](\tau) = \sum_{N\in\mathbb{Z}^g} \exp\left(2\pi i\left\{\frac{1}{2}\,{}^t(N+\epsilon)\tau(N+\epsilon) + {}^t(N+\epsilon)\epsilon'\right\}\right). \tag{2.34}$$

3. Honda-Tate Theorem

Let k be a field, A/k an abelian variety of dimension g. Let ℓ be a prime such that $(\ell, \mathrm{char}(k)) = 1$ if $\mathrm{char}(k) \neq 0$.

Define the ℓ-adic *Tate module*

$$T_\ell(A) = \varprojlim A[\ell^n](\overline{k}). \tag{3.1}$$

An element of $T_\ell(A)$ is thus a sequence $(x_n)_{n\geq 0}$, where $x_0 = 0$, and $\ell x_{n+1} = x_n$. It has a natural $\mathrm{Gal}(\overline{k}/k)$ action.

Exercise 3.1. Let Γ be an abelian group, then we may define $T_\ell(\Gamma)$ in the same fashion: $T_\ell(\Gamma) = \varprojlim \Gamma[\ell^n]$. It may again be described as sequences $(x_n)_{n\geq 0}$, where $x_n \in \Gamma$, $x_0 = 0$, and $\ell x_{n+1} = x_n$. The addition is by component-wise addition and the zero element is the sequence consisting of $x_n = 0$ for all n.

Show that $T_\ell(\Gamma)$ is a torsion-free $\mathbb{Z}_\ell$ module and that $Aut(\Gamma)$ acts as automorphisms on $T_\ell(\Gamma)$. (N.B. When going back to the case of abelian varieties note that $\Gamma = A(\overline{k})$ and not A. In particular, $Gal(\overline{k}/k)$ acts on $A(\overline{k})$, and it does not act as automorphisms of the variety A! The representation $Gal(\overline{k}/k) \longrightarrow \mathrm{Aut}(T_\ell(A))$ is called the ℓ-adic representation).

An important example is taking Γ to be the group of roots of unity of ℓ-power order in k. We shall denote $T_\ell(\Gamma)$ in this case by $\mathbb{Z}_\ell(1)$.

If Γ and Δ are G-modules and

$$\Gamma \times \Gamma \longrightarrow \Delta \tag{3.2}$$

is a G-equivariant alternating pairing, we have an induced G-equivariant pairing

$$T_\ell(\Gamma) \times T_\ell(\Gamma) \longrightarrow T_\ell(\Delta), \tag{3.3}$$

and hence a G-equivariant homomorphism

$$\bigwedge_{\mathbb{Z}_\ell}^2 T_\ell(\Gamma) \longrightarrow T_\ell(\Delta). \tag{3.4}$$

That is, the following diagram commutes:

$$\begin{array}{ccc} \bigwedge^2 T_\ell(\Gamma) & \longrightarrow & T_\ell(\Delta) \\ \downarrow{\scriptstyle \wedge^2 g} & & \downarrow{\scriptstyle g} \\ \bigwedge^2 T_\ell(\Gamma) & \longrightarrow & T_\ell(\Delta) \end{array} . \tag{3.5}$$

Use this to show that for an elliptic curve E over a totally real number field k, taking

$$\Gamma = E(\overline{k}), \quad \Delta = \varprojlim_n \mu_{\ell^n} \tag{3.6}$$

and the Weil pairing, we get that the ℓ-adic representation is *odd*. That is, complex conjugation has determinant -1.

We come back to the case of abelian varieties. If A and B are two abelian varieties over k, then we have a natural map commuting with the Galois group action,

$$\mathrm{Hom}_k(A,B) \otimes_{\mathbb{Z}} \mathbb{Z}_\ell \longrightarrow \mathrm{Hom}_{G_k}(T_\ell(A), T_\ell(B)). \tag{3.7}$$

We remark that the module $T_\ell(A)$ is a free $\mathbb{Z}_\ell$-module of rank $2g$. The map in (3.7) is *always* injective with torsion free cokernel. When k is a finite field (Tate [**112**]), or a function field over a finite field (Zarhin [**125**], [**126**]) or a number field (Faltings [**32**]) or a finitely generated field over $\mathbb{Q}$ (Faltings [**32**]) then it is an isomorphism. There are obvious cases when this is not so.

Let A/k be an abelian variety, $k = \mathbb{F}_{p^r}$. We refer the reader to Appendix A for definition and discussion of the Frobenius and Verschiebung morphisms, and to properties of group schemes. We define $\Phi = \mathrm{Fr}_A^r = F_{A^{(p^{r-1})}} \circ \cdots \mathrm{Fr}_{A^{(p)}} \circ \mathrm{Fr}_A$. Then

$$\Phi : A \longrightarrow A^{(p^r)} = A \tag{3.8}$$

is a homomorphism of degree $(p^g)^r = p^{gr}$ and $\ker(\Phi)$ is a finite connected group scheme of A of order p^{gr}. Using Ver we see that

$$\mathrm{Ker}(\Phi) \subset A[p^r]. \tag{3.9}$$

Moreover, since $\mathrm{Ker}(\Phi)$ is a connected group scheme,

$$\mathrm{Ker}(\Phi)(\overline{k}) = \{0\}. \tag{3.10}$$

This may also be seen directly: $x \in A(\overline{k})$ implies that there exists an ℓ such that $\Phi^\ell(x) = x$. Incidentally, since $A[p]$ is a group scheme of order p^{2g}, this shows again that $|A[p](\overline{k})| \leq p^g$.

For any prime ℓ, including $\ell = p$, we let:

$$T_\ell(A) = \varprojlim_{\longleftarrow} A[\ell^n]. \tag{3.11}$$

Whether $\ell = p$ or not, we get an endomorphism:

$$\Phi \in \mathrm{End}_k(T_\ell(A)). \tag{3.12}$$

Let Δ be the characteristic polynomial of Φ (for $\ell \neq p$). One can prove that Δ has coefficients in $\mathbb{Z}$, is independent of ℓ and has degree $2g$.

THEOREM 3.2. (Honda-Tate)

1. *Let A be a $\mathbb{F}_q$-simple abelian variety. Then $E = \mathrm{End}_{\mathbb{F}_q}(A) \otimes_{\mathbb{Z}} \mathbb{Q}$ is a division algebra with center $\mathbb{Q}[\pi]$, $\pi = \mathrm{Fr}_q = (\mathrm{Fr}_p)^r, q = p^r$. The number π is an algebraic integer such that for every embedding $\mathbb{Q}[\pi] \hookrightarrow \mathbb{C}$, we have $|\pi| = \sqrt{q}$ (We call such an algebraic integer π a q-Weil number).*

2. *For any prime ν of $\mathbb{Q}[\pi]$,*

$$inv_\nu(E) = \begin{cases} \frac{1}{2} & \nu \text{ real} \\ 0 & \nu \text{ complex} \\ \frac{ord_\nu(\pi)}{ord_\nu(q)} [\mathbb{Q}[\pi]_\nu : \mathbb{Q}_p] & \nu | p \\ 0 & \text{otherwise} \end{cases} \tag{3.13}$$

Moreover, $\dim A = \frac{1}{2}[\mathbb{Q}[\pi] : \mathbb{Q}] \cdot \sqrt{[E : \mathbb{Q}[\pi]]}$ *and* $e = \sqrt{[E : \mathbb{Q}[\pi]]}$ *is the lowest common denominator of the* $inv_v(E)$*'s. The characteristic polynomial of* π *is* $m(T)^e$*, where* $m(T)$ *is the minimal polynomial of* π *over* $\mathbb{Q}$*.*

3. *Two simple abelian varieties A and A' over $\mathbb{F}_q$ are isogenous over $\mathbb{F}_q$ iff there exists an isomorphism $\phi : \mathbb{Q}[\pi_A] \longrightarrow \mathbb{Q}[\pi'_A]$ such that $\phi(\pi_A) = \pi'_A$.*
4. *Every q-Weil number comes from a $\mathbb{F}_q$-simple abelian variety.*

REMARK 3.3. 1. In Part 1 of the theorem, the abelian variety does not need to be $\overline{\mathbb{F}_q}$-simple. For example, put $\pi = p$. Hence $q = p^2$, and

$$inv_\nu(E) = \begin{cases} \frac{1}{2} & \nu \text{ real} \\ \frac{1}{2} & \nu = p \\ 0 & \text{otherwise} \end{cases} . \tag{3.14}$$

We infer that π corresponds to an elliptic curve X over $\mathbb{F}_{p^2}$, and $E = \mathrm{End}_{\mathbb{F}_{p^2}}(X) \cong B_{p,\infty}$, the "unique" quaternion algebra over $\mathbb{Q}$ ramified at p and ∞. This is a supersingular elliptic curve.

Now do the calculation with $\pi = \sqrt{p}$, so $q = p$ and $\mathbb{Q}[\pi] = \mathbb{Q}[\sqrt{p}]$.

$$inv_\nu(E) = \begin{cases} \frac{1}{2} & \nu \text{ real} \\ 0 & \text{otherwise} \end{cases} . \tag{3.15}$$

The quaternion algebra E over $\mathbb{Q}[\sqrt{p}]$ is split at all finite places and ramified at the two real places. Then, the corresponding abelian variety A is 2 dimensional, and A is simple abelian surface over $\mathbb{F}_p$. But A it is not simple over $\mathbb{F}_{p^2}$, because over $\mathbb{F}_{p^2}$ we get the previous computation leading to the elliptic curve X. In fact, A is isogenous to a product $X_1 \times X_2$, where the X_i are elliptic curves. Moreover, since $\pi^2 = p$ induces the Frobenius morphism of each $X_i/\mathbb{F}_{p^2}$, we see that X_1 and X_2 are both $\mathbb{F}_{p^2}$-isogenous to X above.

2. There is a bijection between $\mathbb{F}_q$-isogeny classes of simple abelian varieties over $\mathbb{F}_q$ and characteristic polynomials of q-Weil numbers.
3. Take a supersingular elliptic curve over $\mathbb{F}_q$. Enlarging q, we have $E = B_{p,\infty}$. Then $\pi \in \mathbb{Q}$ and replacing q by q^2, we get $\pi = q$. Conclusion: Every two supersingular elliptic curves are isogenous over an algebraically closed field.
4. Examples of q-Weil numbers giving elliptic curves: Let π be a solution to $X^2 - \beta X + q = 0$ for $|\beta| < 2\sqrt{q}, \beta \in \mathbb{Z}$. Then $\pi = \frac{\beta \pm \sqrt{\beta^2 - 4q}}{2}$ is totally imaginary, $|\pi| = \sqrt{q}$, hence $q = \pi\overline{\pi}$. If π is not associated to $\overline{\pi}$, then p is split in $\mathbb{Q}[\pi]$.

See [**111**],[**74**], [**119**], [**123**], [**100**] for more on this fascinating subject.

EXERCISE 3.4. Estimate the number of isogeny classes of elliptic curves over $\mathbb{F}_q$.

EXERCISE 3.5. Fix q, $(q,2)=1$, and an elliptic curve $E: y^2 = x^3 + ax + b$ over $\mathbb{F}_q$. For every $d \in \mathbb{F}_q$ let E_d be the quadratic twist $dy^2 = x^3 + ax + b$ of E. Among all quadratic twists, how many isomorphism classes are there? how many isogeny classes?

4. Ordinary Abelian Varieties and Serre-Tate Coordinates

4.1. Ordinary abelian varieties. Let k be an algebraically closed field of characteristic $p > 0$. An abelian variety A over k is *ordinary* if

$$|A[p](k)| = p^g. \tag{4.1}$$

As explained in Section 1, this is the maximal power of p possible. The property of being ordinary is stable under isogeny. If $A/\mathbb{F}_q$ is an abelian variety and Δ is the characteristic polynomial of Fr_q, then we will see later the A is ordinary iff the Newton polygon has two slopes, 0 and 1. N.B.: Ordinary is an assertion about the Newton polygon of Δ, while the isogeny class is an assertion about Δ itself, and hence stronger. In fact

EXERCISE 4.1. Use Honda-Tate, and in particular Remark (3.3)(4), to show that there are infinitely many isogeny classes of ordinary elliptic curves over $\overline{\mathbb{F}_p}$.

REMARK 4.2. Ordinary abelian varieties are of paramount importance because they are almost always dense subsets of moduli spaces ([**120**]), ordinary is an open property, and ordinary abelian varieties can be studied most effectively using Serre-Tate theory.

Let $\mathcal{C}_k$ be the category with:

1. **Objects** consisting of triples $(R, \mathfrak{m}_R, \phi)$, where R is a local Artinian ring, $\mathfrak{m}_R$ is its maximal ideal, and $\phi : R/\mathfrak{m}_R \xrightarrow{\sim} k$.
2. **Morphisms**: $\mathrm{Mor}(R_1, R_2)$ are local homomorphisms $\psi : R_1 \longrightarrow R_2$ inducing the identity on k. Recall that a local homomorphism of local rings is a homomorphism satisfying $\mathfrak{m}_{R_1} = \phi^{-1}(\mathfrak{m}_{R_2})$.

Consider the functor $\mathcal{F}$ of local deformations of A/k:

$$\mathcal{C}_k \longrightarrow \mathbf{Sets}, \tag{4.2}$$

given by

$$\mathcal{F}(R) = \{\mathcal{A}/R \text{ an abelian scheme with an identification } \Phi : \mathcal{A} \otimes_R k \xrightarrow{\sim} A, \text{ up to isomorphisms inducing the identity on } A\}. \tag{4.3}$$

By fundamental results of Grothendieck, Mumford and Schlessinger (see [**91**] for discussion and references), there exists a complete local noetherian ring $\mathcal{O} = \widehat{\mathfrak{M}_{A/k}}$ [3] with residue field k such that the functor of points h of $\mathrm{Spec}(\mathcal{O})$ pro-represent $\mathcal{F}$.

[3] This notation is intended to suggest that $\mathcal{O}$ is the completion of the local ring of the point corresponding to A/k in the moduli space $\mathcal{M}$ of abelian varieties, if such existed. In fact, if one rigidifies the situation by introducing more structure into the deformation problem (e.g., polarization, real multiplication) then the corresponding complete local ring that pro-represents the new functor (it would be a quotient of $\mathcal{O}$ by some ideal) is the completion of the local ring of the moduli space $\mathcal{M}$ parameterizing such an object at the point corresponding to A + the extra data.

As a *General Principle*, if one has a fine moduli space $\mathfrak{M}$ representing a functor $\mathcal{F}$, and $x \in \mathfrak{M}(k)$, then $\widehat{\mathfrak{M}}_x$ – the completion of the local ring of $\mathfrak{M}$ at x – pro-represents the local deformation functor $\mathcal{F}|_{\mathcal{C}_k}$ with the initial data given by the object parameterized by x.

That is

$$h(R) = \mathrm{Mor}_{\mathrm{Spec}(k)}(\mathrm{Spec}\, R, \mathrm{Spec}\, \mathcal{O}) = \mathrm{Hom}(\mathcal{O}, R) = \mathcal{F}(R), \quad \forall R \in \mathcal{C}_k, \tag{4.4}$$

where the ring homomorphisms are local homomorphisms. We shall discuss those notions in more detail in Chapter 6. In particular, we shall state precisely a theorem of Serre and Tate that asserts that the deformations of an abelian variety A over an algebraically closed field k of characteristic p are the same as the deformations of its p-divisible group $\varinjlim A[p^n]$. For ordinary abelian varieties over k the p-divisible group is composed of the local part, which is the completion of a split torus, and the étale part which is a constant group scheme isomorphic to $(\underline{\mathbb{Q}_p/\mathbb{Z}_p})^g$. Rigidity for tori implies that the deformations are all given by extension classes of these two groups. It is thus possible to describe those in a very explicit manner. This is done in [**61**]. We describe the final result below.

4.2. Serre-Tate coordinates. Let A be an abelian variety over an algebraically closed field k of characteristic p.

THEOREM 4.3. (Serre-Tate Coordinates)

1. *We have an isomorphism of functors on the category $\mathcal{C}_k$:*

$$\widehat{\mathfrak{M}}_{A/k}(-) \cong \mathrm{Hom}_{\mathbb{Z}_p}((T_pA)(k) \otimes_{\mathbb{Z}_p} (T_pA^\vee)(k), \widehat{\mathbb{G}_m}(-)), \quad R \in \mathcal{C}_k. \tag{4.5}$$

2. *Let $R \in \mathcal{C}_k$, and let $\mathcal{A}/R$, $\mathcal{B}/R$ be deformations of two ordinary abelian varieties A/k, B/k, respectively, with corresponding bilinear forms $q_{\mathcal{A}}$, $q_{\mathcal{B}}$. A homomorphism $f : A \longrightarrow B$ extends to homomorphism $\mathbf{f} : \mathcal{A} \longrightarrow \mathcal{B}$ if and only if*

$$q_{\mathcal{A}}(x, f^\vee(y)) = q_{\mathcal{B}}(f(x), y), \quad \forall x \in T_pA(k), y \in T_pB^\vee(k). \tag{4.6}$$

We recall that if A is ordinary, then $T_pA(k), T_pA^\vee(k)$ are free $\mathbb{Z}_p$-modules of rank g, and that $\widehat{\mathbb{G}_m}$ is defined by

$$\widehat{\mathbb{G}_m} = \varinjlim \mu_{p^n}. \tag{4.7}$$

EXERCISE 4.4. Prove that for $R \in \mathcal{C}_k$, $\widehat{\mathbb{G}_m}(R) = 1 + \mathfrak{m}_R$. Thus (4.5) says that for all $R \in \mathcal{C}_k$:

$$\mathcal{F}(R) = \widehat{\mathfrak{M}}_{A/k}(R) = \mathrm{Hom}_{\mathbb{Z}_p}(T_pA(k) \otimes_{\mathbb{Z}_p} T_pA^\vee(k), 1 + \mathfrak{m}_R). \tag{4.8}$$

Choosing coordinates for $T_pA(k), T_pA^\vee(k)$ as $\mathbb{Z}_p$ modules, we find

$$\widehat{\mathfrak{M}}_{A/k} \cong \mathbb{Z}_p^{g^2} \otimes_{\mathbb{Z}_p} \widehat{\mathbb{G}_m} \cong \widehat{\mathbb{G}_m}^{g^2}. \tag{4.9}$$

More importantly, for every $R \in \mathcal{C}_k$, the deformations of A over R, $\widehat{\mathfrak{M}}_{A/k}(R) = \mathcal{F}(R)$, correspond to $\mathbb{Z}_p$-linear maps:

$$\mathcal{A}/R \longleftrightarrow <,>_{\mathcal{A}} : T_pA(k) \times T_pA^\vee(k) \longrightarrow 1 + \mathfrak{m}_R. \tag{4.10}$$

A remarkable consequence of the existence of the Serre-Tate coordinates is the existence a natural group structure on $\mathfrak{M}_{A/k}$. This group law is simply given by "addition" of bilinear forms (N.B.: The bilinear forms are with values in multiplicative group. Thus the sum of bilinear forms is obtain by multiplying their values). In particular, it has a distinguished element, the identity of the group which is just the trivial bilinear pairing $<,> \equiv 1$.

DEFINITION 4.5. For every R, we let $\mathcal{A}^{can}/R$ be the deformation corresponding to the pairing $<,>\equiv 1$. It is called the *canonical lift* of A to R.

By taking limits, we can also get a canonical lift to $W(k)$ (this uses results of Grothendieck on algebraization of formal schemes and uses implicitly the existence of a polarization on the "limit" object "$\varinjlim \mathcal{A}^{can}/(W(k)/p^nW(k))$").

One sees from the criterion furnished by (4.6) that $\mathrm{End}_k(A)$ lifts to $\mathcal{A}^{can}/k$. In general, a deformation $\mathcal{A}/W(k)$ such that $\mathrm{End}_k(A)$ lifts to $\mathcal{A}/W(k)$ corresponds to a $W(k)^\times$-valued pairing satisfying

$$< x, f^\vee(y) >=< f(x), y > \forall f \in \mathrm{End}(A), x \in T_pA(k), y \in T_pA^\vee(k). \tag{4.11}$$

If $k = \overline{\mathbb{F}_p}$, then this property *characterizes* the canonical lift. Essentially, because A is defined over some $\mathbb{F}_q$, and then the endomorphisms of A given by $f = \mathrm{Fr}_q$, $f^\vee = \mathrm{Ver}_q$, behave very differently. The action of f and $f^\vee$ on the group schemes μ_{q^r} and $\underline{\mathbb{Z}/q^r\mathbb{Z}}$ are given by the following table:

(4.12)

	μ_{q^r}	$\underline{\mathbb{Z}/q^r\mathbb{Z}}$
f	raising to q power; $\mathrm{Ker}(f) = \mu_q$	id
$f^\vee$	id	multiplying by q; $\mathrm{Ker}(f^\vee) = \underline{q^{r-1}\mathbb{Z}/q^r\mathbb{Z}}$

We note that over k, $A[q^r] \cong \mu_{q^r}^g \oplus (\underline{\mathbb{Z}/q^r\mathbb{Z}})^g$. Assume that $\mathcal{A}/W(k)$ is a deformation with a bilinear form $<,>$ such that every endomorphism of A lifts to $\mathcal{A}$. We want to show that $\mathcal{A} = \mathcal{A}^{can}$. Let U_1 denote the units of $W(k)$ that are congruent to 1 modulo p. It is enough to show that for every r the induced pairing

$$<,>: T_pA(k)/q^rT_pA(k) \times T_pA^\vee(k)/q^rT_pA^\vee(k) \longrightarrow U_1/U_1^{q^r}, \tag{4.13}$$

is trivial. Apply f^r. Since f^r is an endomorphism of $\mathcal{A}$, we must have

$$< x, (f^\vee)^r y >=< f^r(x), y >, \ \ \forall (x, y) \in T_pA(k) \times T_pA^\vee(k). \tag{4.14}$$

However, on $T_pA(k)/q^rT_pA(k) \cong A[q^r](k) \cong (\underline{\mathbb{Z}/q^r\mathbb{Z}})^g$ the morphism f^r is the identity, while on $T_pA^\vee(k)/q^rT_pA^\vee(k) \cong A^\vee[q^r](k) \cong (\underline{\mathbb{Z}/q^r\mathbb{Z}})^g$, the morphism $(f^\vee)^r$ is zero. This proves that (4.13) is trivial.

We further remark, that if λ is a principal polarization of A, and we use it to identify T_pA with $T_pA^\vee$, then the deformations of (A, λ) (as a principally polarized abelian variety) are given by the symmetric elements of

$$\mathrm{Hom}_{\mathbb{Z}_p}(T_pA(k) \otimes_{\mathbb{Z}_p} T_pA^\vee(k), \widehat{\mathbb{G}_m}(-)). \tag{4.15}$$

EXAMPLE 4.6. If $\mathfrak{M}$ is the moduli space of principally polarized abelian varieties (with level n structure, $n \geq 3$), $x \in \mathfrak{M}(k)$, then

$$\widehat{\mathfrak{M}}_x = \widehat{\mathfrak{M}}_{(A,\lambda)/k} \cong \widehat{\mathbb{G}_m}^{\frac{g(g+1)}{2}}, \tag{4.16}$$

where $\mathfrak{M}_x$ is the formal spectrum of the completion of the local ring of x in $\mathfrak{M}$, where $\widehat{\mathfrak{M}}_{(A,\lambda)/k}$ is the complete local ring that pro-represents the functor of deformations

of (A,λ) over $\mathcal{C}_k$, and where the last isomorphism is valid for A ordinary. Thus, at least at ordinary points, we know that $\mathfrak{M}$ is non-singular of dimension $\frac{g(g+1)}{2}$ (though this is true at every point).

Following Chai-Norman [**11**] we give another application of the Serre-Tate coordinates.

DEFINITION 4.7. A $\Gamma_0(p)$ *level structure* is a pair of two principally polarized abelian variety (A,λ_A) and (B,λ_B) of dimension g, and an isogeny:

$$(A,\lambda_A) \xrightarrow{f} (B,\lambda_B), \tag{4.17}$$

such that $f^*\lambda_B = p\lambda_A$.

Put $V = T_pA$, $W = T_pB$. Let us compute the structure of the moduli space of $\Gamma_0(p)$ level structure at a formal neighborhood of an ordinary point. The local deformation functor $\mathcal{C}_k \longrightarrow$ **Sets** is given by:

$$\begin{aligned} R \longmapsto \{ &\text{ symmetric pairings} \\ &<,>_V\colon V\otimes V \longrightarrow \widehat{\mathbb{G}_m}(R),\ \ <,>_W\colon W\otimes W \longrightarrow \widehat{\mathbb{G}_m}(R), \\ &\qquad \text{such that } \ < v, f^\vee(w) >_V =< f(v), w)_W, \text{ for } v\in V, w\in W\}. \end{aligned} \tag{4.18}$$

By the Elementary Divisors Theorem, we may choose bases $(v_1,\dots,v_g)$ for V, and $(w_1,\dots,w_g)$ to W, such that

$$f(v_1) = w_1,\dots,f(v_a) = w_a;\ \ f(v_{a+1}) = pw_{a+1},\dots,f(v_g) = pw_g, \tag{4.19}$$

and hence

$$f^\vee(w_1) = pv_1,\cdots,f^\vee(w_a) = pv_a;\ \ f^\vee(w_{a+1}) = v_{a+1},\cdots,f^\vee(w_g) = v_g. \tag{4.20}$$

We have a $\frac{g(g+1)}{2}$-dimensional space of pairings such that:

$$< v_i, f^\vee(w_j) >_V =< f(v_i), w_j >_W, \quad \forall i,j. \tag{4.21}$$

More precisely, the conditions

$$< v_i, pv_j >_V =< w_i, w_j >_W, \quad 1\le i\le a,\ 1\le j\le a, \tag{4.22}$$

give an $\frac{a(a+1)}{2}$-dimensional space; the conditions

$$< v_i, v_j >_V =< w_i, w_j >_W, \quad 1\le i\le a,\ a+1\le j\le g, \tag{4.23}$$

give an $a(g-a)$-dimensional space; the conditions

$$< v_i, v_j >_V =< pw_i, w_j >_W, \quad a+1\le i\le g,\ a+1\le j\le g, \tag{4.24}$$

give a $\frac{(g-a)(g-a+1)}{2}$-dimensional space. Together we get $a(a+1)/2 + a(g-a) + (g-a)(g-a+1)/2 = g(g+1)/2$ dimensions. Therefore, the following theorem is proved.

THEOREM 4.8. (Chai-Norman) *The coarse moduli space $\mathcal{M}$ for $\Gamma_0(p)$-level structure at an ordinary point $f : (A,\lambda_A) \longrightarrow (B,\lambda_B)$ satisfies*

$$\widehat{\mathfrak{M}}_{(f:A\longrightarrow B)} \cong \mathbb{Z}_p^{\frac{g(g+1)}{2}} \otimes_{\mathbb{Z}_p} \widehat{\mathbb{G}_m}. \tag{4.25}$$

5. Abelian Varieties with Real Multiplication over a General Field

Let k be a field. Let L be a totally real field of degree g with discriminant d_L.

DEFINITION 5.1. An abelian variety with real multiplication by $\mathcal{O}_L$ (or, an "abelian variety with RM" for short) is an abelian variety A over k together with an embedding of rings

$$\iota : \mathcal{O}_L \hookrightarrow \mathrm{End}_k(A), \tag{5.1}$$

such that the Deligne-Pappas condition holds:

(DP) $A \otimes_{\mathcal{O}_L} \mathcal{M}_A \cong A^\vee$.

A few remarks are in order. First, since $\mathcal{O}_L$ acts on A it also acts on $A^\vee$ by duality. One defines then the symmetric $\mathcal{O}_L$-linear homomorphisms

$$\mathcal{M}_A = \{\lambda : A \longrightarrow A^\vee : \lambda \circ \iota(r) = \iota(r)^\vee \circ \lambda,\ \forall r \in \mathcal{O}_L\}. \tag{5.2}$$

The abelian variety $A \otimes_{\mathcal{O}_L} \mathcal{M}_A$ and the homomorphism $A \otimes_{\mathcal{O}_L} \mathcal{M}_A \cong A^\vee$ are all determined uniquely by their behavior on k-algebras R:

$$(A \otimes_{\mathcal{O}_L} \mathcal{M}_A)(R) = A(R) \otimes_{\mathcal{O}_L} \mathcal{M}_A;\ \ x \otimes \lambda \mapsto \lambda(x). \tag{5.3}$$

Secondly, one can show that if A satisfies the **(DP)** condition, so does $A^\vee$. The module $\mathcal{M}_A$ is a projective $\mathcal{O}_L$-module of rank one endowed with a natural notion of positivity determined by the cone of polarizations in $\mathcal{M}_A$. C.f. Chapter 2, Section 2.2.

Thirdly, if $(\mathrm{char}(k), d_L) = 1$ or $\mathrm{char}(k) = 0$ then the **(DP)** condition is equivalent to Rapoport's condition :

(R) $\mathfrak{t}^*_{A/k}$ is a free $\mathcal{O}_L \otimes_{\mathbb{Z}} k$-module of rank 1, and $\mathcal{M}_A$ is a projective $\mathcal{O}_L$-module of rank 1.

This condition is heuristically easier to understand. We remark that for abelian varieties in characteristic zero, or for ordinary abelian varieties condition **(R)** holds automatically, provided they have an $\mathcal{O}_L$-linear polarization. See Chapter 2 Corollary 2.5 and Corollary 6.4 below.

Let us engage again in Kodaira-Spencer heuristics:

The moduli space of complex abelian varieties with RM is g-dimensional. If a reasonable moduli scheme, say over $\mathrm{Spec}(\mathbb{Z})$, for abelian varieties with RM exists at all, we expect it then to be of relative dimension g. The Kodaira-Spencer method (see Chapter 1 Remark 6.8) allows us to compute the tangent space to the moduli problem. It is given by

$$\mathrm{Hom}_{\mathcal{O}_L}(\mathfrak{t}^*_A, \mathfrak{t}_{A^\vee}). \tag{5.4}$$

If condition **(R)** holds this is a g-dimensional space. Thus condition **(R)** is natural to impose. However, it turns out that the moduli scheme is then not proper over $\mathrm{Spec}(\mathbb{Z})$, not only because of lack of cusps (that are necessary because degenerations of abelian varieties with RM to semi-abelian varieties with RM exist), but because of families of abelian varieties with RM satisfying **(R)** degenerating into abelian varieties with $\mathcal{O}_L$-action not satisfying **(R)**. Deligne and Pappas discovered [**21**] that abelian varieties with RM satisfying **(R)** satisfy **(DP)** and, moreover, the property of **(DP)** is preserved under degenerations.

It also turns out, see [**21**] (see also [**3**]), that the moduli scheme is singular at all fibers dividing d_L, the singular locus is of codimension 2 and consists precisely of all points where condition **(R)** fails.

LEMMA 5.2. *The Condition* (**R**) *always implies Condition* (**DP**). *If* $(p, d_L) = 1$*, the conditions are equivalent.*

PROOF. **To be supplied.** □

Examples of abelian varieties with real multiplication can be constructed in characteristic zero by means of the complex uniformization given in Chapter 2 Section 2.2. One can then obtain examples in characteristic p by reducing modulo a suitable prime ideal.

Since the moduli space of abelian varieties with RM is smooth over $\mathbb{Z}[d_L^{-1}]$, every abelian variety in characteristic p, $(p, d_L) = 1$, can be lifted to characteristic zero. The same holds for characteristic p such that $p|d_L$, but one needs a more elaborate argument, due essentially to Mumford (see [**91**, p. 249]). It is a general principle saying, roughly, that if your object (in this case an abelian variety with RM) is a special fibre on a nice enough family, over an equi-characteristic integral scheme, of objects of the same kind. And if the generic fibre of such this family lifts to characteristic zero, then the special fibre lifts as well.

It is enough therefore, given an abelian variety with RM in characteristic p, represented by a point x in the moduli space, to construct a smooth curve in the moduli space modulo p that passes through x and whose generic fibre is non-singular point of the ambient moduli space; namely, that the relative cotangent is generically free. At least for $g = 2$ such a curve may be constructed using a Moret-Bailly family (See [**3**]). In general, such a curve exists because the moduli space is quasi-projective and the singularities are of codimension 2.

Another way to produce examples is to use the Honda-Tate Theorem (Theorem 3.2) to construct first isogeny classes having multiplication by some order in $\mathcal{O}_L$, and then finding particular representatives for this isogeny class that have multiplication by the whole of $\mathcal{O}_L$.

6. Irreducibility of the Moduli Space of μ_{p^∞}-level Structure

Let L be a totally real field of degree $g = [L : \mathbb{Q}]$, with ring of integers $\mathcal{O}_L$, different $\mathcal{D}_L$, and discriminant d_L . In this section, we follow Ribet's proof of the irreducibility of the moduli space of abelian varieties in characteristic p with μ_p-level structure when $(p, d_L) = 1$. The argument is essentially that of monodromy.

We first give the scheme theoretic analogue of our definition of abelian varieties with RM generalizing the notion of abelian variety with real multiplication of Definition 5.1.

DEFINITION 6.1. Let S be a scheme in which d_L is invertible. An *abelian scheme with real multiplication* by $\mathcal{O}_L$ is a couple $(A/S, \iota)$, where A/S is an abelian scheme and

$$\iota : \mathcal{O}_L \hookrightarrow \mathrm{End}(A/S) \tag{6.1}$$

is a ring injection such that $\mathfrak{t}_{A/S}$, the relative tangent sheaf (also denoted by $\mathrm{Lie}(A/S)$), is a locally free $\mathcal{O}_L \otimes_{\mathbb{Z}} \mathcal{O}_S$-module of rank 1, and such that the $\mathcal{O}_L$-module of symmetric $\mathcal{O}_L$-linear homomorphisms to the dual abelian scheme is a projective rank 1 module in the étale topology.

REMARK 6.2. If we let $A^\vee/S$ be the dual abelian scheme, and define:

$$\iota^\vee : \mathcal{O}_L \longrightarrow \mathrm{End}(A^\vee/S), \quad \iota^\vee(m) = (\iota(m))^\vee. \tag{6.2}$$

Then $(A^\vee/S, \iota^\vee)$ is again a abelian scheme with real multiplication.

Let us restrict our attention to the case of an algebraically closed field k of characteristic p not dividing the discriminant. Let A/k be an ordinary abelian variety of dimension g with real multiplication by $\mathcal{O}_L$. Consider the Tate module $T_pA(k)$ ($\cong \mathbb{Z}_p^g$) as a module over $\mathbb{Z}_p \otimes \mathcal{O}_L$. Note that $\mathbb{Z}_p \otimes \mathcal{O}_L$ acts faithfully on $T_pA(k)$ because $\mathbb{Z}_p/p^n\mathbb{Z} \otimes \mathcal{O}_L$ acts faithfully on $A[p^n](k)$:

One can show that A has a polarization of degree prime to p (see the proof of Lemma 5.2), hence $A[p]$ is a self-dual group scheme, since this polarization on A induces an isomorphism of $A[p]$ with $A^\vee[p] = A[p]^\vee$. So,

$$A[p^n](k) \cong (\mu_{p^n})^g \oplus (\underline{\mathbb{Z}/p^n\mathbb{Z}})^g \cong \mu_{p^n} \otimes \mathcal{O}_L \oplus (\underline{\mathbb{Z}/p^n\mathbb{Z}}) \otimes \mathcal{O}_L, \tag{6.3}$$

and the faithfulness of the action of $\mathbb{Z}_p/p^n\mathbb{Z} \otimes \mathcal{O}_L$ on $\mu_{p^n} \otimes \mathcal{O}_L$ is equivalent to the faithfulness on $(\mu_{p^n})^g$. But,

$$\mathrm{Lie}(\mu_p \otimes \mathcal{O}_L) = \mathrm{Lie}(A/k) = \mathfrak{t}_{A/k}. \tag{6.4}$$

Therefore $\mathbb{Z}_p/p^n\mathbb{Z} \otimes \mathcal{O}_L$ acts faithfully on $\mu_{p^n} \otimes \mathcal{O}_L$ and by duality of $\mathcal{O}_L$-groups also on $\mathbb{Z}/p^n\mathbb{Z} \otimes \mathcal{O}_L$.

COROLLARY 6.3. *The $\mathbb{Z}_p \otimes \mathcal{O}_L$-module $T_pA(k)$ is free of rank* 1.

COROLLARY 6.4. *An ordinary abelian variety with RM satisfies condition* **(R)**.

What are the deformations of $(A/k, \iota)$?

Apply Serre-Tate coordinates: for $R \in \mathcal{C}_k$, we look for the pairings

$$<,>: T_pA(k) \otimes T_pA^\vee(k) \longrightarrow \widehat{\mathbb{G}_m}(R), \tag{6.5}$$

such that:

$$< x, my >=< mx, y >, \quad \forall x \in T_pA(k),\ y \in T_pA^\vee(k),\ m \in \mathcal{O}_L. \tag{6.6}$$

Fix a generator y of $T_pA^\vee(k)$ as an $\mathbb{Z}_p \otimes \mathcal{O}_L$ module. Define

$$\phi_y : T_pA(k) \longrightarrow \widehat{\mathbb{G}_m}, \ \ \phi_y(x) =< x, y > . \tag{6.7}$$

Note that $\phi_y(mx) =< x, my >$, and hence ϕ_y determines the pairing. Conversely, given a homomorphism $\phi : T_pA(k) \longrightarrow \widehat{\mathbb{G}_m}$ define a pairing by $< x, my >= \phi(mx)$. One checks that this is a well defined $\mathcal{O}_L$-bilinear pairing $<,>: T_pA(k) \otimes T_pA^\vee(k) \longrightarrow \widehat{\mathbb{G}_m}(R)$. Thus the $\mathcal{O}_L$-pairings are identified with $\mathrm{Hom}(T_pA(k), \widehat{\mathbb{G}_m})$, and one concludes:

PROPOSITION 6.5. *There is an isomorphism*

$$\widehat{\mathcal{M}_{(A/k,\iota)}} \cong \widehat{\mathbb{G}_m}^g. \tag{6.8}$$

DEFINITION 6.6. A *full level n structure* on a abelian scheme with real multiplication $(A/S, \iota)$ is an isomorphism of constant group schemes with $\mathcal{O}_L$-action

$$\alpha : (\mathcal{O}_L/n\mathcal{O}_L)^2 \xrightarrow{\sim} A[n]. \tag{6.9}$$

(Note that we do not take necessarily a symplectic level structure).

THEOREM 6.7. (Rapoport [**97**, Lemme 1.23, p.267]) *If $n \geq 3$ the functor over $\mathbb{Z}[\frac{1}{nd_L}]$-schemes:*

$$S \longrightarrow \{(A/S, \iota, \alpha)\} / \cong \tag{6.10}$$

is representable by a smooth morphism $\mathcal{M} \longrightarrow \mathbb{Z}[\frac{1}{nd_L}]$ of relative dimension g.

COROLLARY 6.8. *For $(p, n) = 1$, the components of $\mathcal{M} \otimes \overline{\mathbb{F}_p}$ are the same as the components of $\mathcal{M} \otimes \mathbb{C}$.*

PROOF. Recall Zariski's Main Theorem:

THEOREM 6.9. *Let $f : X \longrightarrow Y$ be a birational projective morphism of noetherian schemes, and assume that Y is normal. Then for every $y \in Y$, $f^{-1}(y)$ is connected. See* [**47**, Corollary 11.4, p. 280, Chapter III].

The Main Theorem yields that the geometric fibers are connected; but since the scheme $\mathcal{M}$ is smooth over its base, these fibers are geometrically regular of equi-dimension g. In fact, the points corresponding to a fixed full level n structure form a component, so the set of components of either $\mathcal{M} \otimes \overline{\mathbb{F}_p}$ or $\mathcal{M} \otimes \mathbb{C}$ is $\mathrm{Isom}(\mathcal{O}_L/n\mathcal{O}_L)^2, A[n]) = \mathrm{Isom}(\mu_n, \mathbb{Z}/n\mathbb{Z})$ (which has cardinality $\phi(n)$.

See [**97**, Théorème 1.28, p. 268, Variante 6.2, p. 325] for more details. □

We may ask how do level structures behave under reduction mod p when the prime p divides the level. After all, a regular scheme over $\mathrm{Spec}(\mathbb{Z}_p)$ with an irreducible generic fibre may very well have a reducible special fibre. E.g., $\mathrm{Spec}(\mathbb{Z}_p[x, y]/(xy - p) \longrightarrow \mathrm{Spec}(\mathbb{Z}_p)$ (Figure ****); or see Kodaira's classification of the special fibre of the Néron model of an elliptic curve (e.g., [**108**, Chapter IV.8]). This is quite a subtle question and to study one such level structure (the μ_{p^r} level structure defined below), just for ordinary points, we would requires Deligne's description of ordinary abelian varieties over finite fields.

Figure 4.

Let A be a g-dimensional abelian variety over $\mathbb{F}_q$, $q = p^r$. Fix an algebraic closure $\overline{\mathbb{F}_q}$ and an embedding $W(\overline{\mathbb{F}_q}) \overset{\phi}{\hookrightarrow} \mathbb{C}$. Assume that A is ordinary; then, via ϕ, we get a canonical lifting $A_{\mathbb{C}}$ of A to $\mathbb{C}$: $A_{\mathbb{C}} := A^{can} \otimes_{W(\overline{\mathbb{F}_q})} \mathbb{C}$. Put

$$T(A) = H_1(A_{\mathbb{C}}, \mathbb{Z}). \tag{6.11}$$

It is a lattice in $\mathbb{C}$ of rank $2g$. Note that $\mathrm{Fr}_q \in \mathrm{End}_{W(\overline{\mathbb{F}_q})}(A) = \mathrm{End}_{\mathbb{C}}(A_{\mathbb{C}})$, hence it defines a linear operator $F \in \mathrm{End}(T(A))$.

THEOREM 6.10. (Deligne [**19**]) *The functor*

$$\{\textit{ordinary abelian varieties over } \mathbb{F}_q\} \longrightarrow \{\textit{free even rank } \mathbb{Z}-\textit{lattices with an endomorphism } F \textit{ such that the conditions } 1,2,3,4 \textit{ hold}\}, \tag{6.12}$$

$$A \longmapsto (T(A), F) \tag{6.13}$$

is an equivalence of categories. The conditions are the following:

1. *F is a semisimple operator all whose eigenvalues have complex absolute value $\sqrt{q}$.*
2. *Half the roots of the characteristic polynomial of F are p-adic units.*
3. *There exists a linear operator V such that $FV = q$.*
4. *There exist free $\mathbb{Z}_p$-modules T'_p, T''_p such that; (a) $T \otimes \mathbb{Z}_p = T'_p \oplus T''_p$, (b) $\mathrm{rank}_{\mathbb{Z}_p} T'_p = \mathrm{rank}_{\mathbb{Z}_p} T''_p$, and (c) $F_{|_{T'_p}}$ is invertible and $F_{|_{T''_p}}$ is divisible by q.*

REMARK 6.11. If the condition 1 holds, then $\{2,3\} \iff 4$.

REMARK 6.12. In fact, starting from $A/\mathbb{F}_q$, $T = T(A)$, then T'_p, T''_p are nonother than $T_pA(\overline{\mathbb{F}_q})$ and $T_pA^\vee(\overline{\mathbb{F}_q})$.

COROLLARY 6.13. *Let $(A/\mathbb{F}_q, \iota)$ be an abelian variety with RM by $\mathcal{O}_L$. We get an induced embedding $T(\iota) : \mathcal{O}_L \longrightarrow \mathrm{End}(T(A))$ such that the image of $\mathcal{O}_L$ commutes with F. Thus,*

$$(A, \iota) \longrightarrow (T(A), F, T(\iota)) \tag{6.14}$$

is an equivalence between ordinary abelian varieties with RM by $\mathcal{O}_L$ over $\mathbb{F}_q$ and free even rank $\mathcal{O}_L$-lattices with an $\mathcal{O}_L$-endomorphism F such that the conditions $1,2,3,4$ hold.

Given $a \in \mathcal{O}_L$ such that $((a),(p)) = 1$, and n such that $a^2 - 4p^n \ll 0$, let

$$T = \mathcal{O}_L[x]/(x^2 - ax + p^n). \tag{6.15}$$

It is a $\mathbb{Z}$-lattice of rank $2g$ and a free $\mathcal{O}_L$-module (of rank 2). Let F be multiplication by x; then $\mathcal{O}_L$ acts on T as ordinary multiplication. Note that $T \otimes \mathbb{Q}$ is a CM field, hence F is semisimple.

$$|x| = |\frac{a \pm \sqrt{a^2 - 4p^n}}{2}| = |\frac{a + \sqrt{a^2 - 4p^n}}{2} \cdot \frac{a - \sqrt{a^2 - 4p^n}}{2}|^{\frac{1}{2}} = \sqrt{p^n}. \tag{6.16}$$

Mod p, x solves the equation $x(x-a) = 0$. This implies that half of the roots of the characteristic polynomial of F are p-adic units. Take $V = -(x-a)$ then $FV = p^n$. **Conclusion:** The conditions imposed in Corollary 6.13 on (T, F, ι) are satisfied and (T, F, ι) corresponds to an ordinary abelian variety A of dimension g with RM by $\mathcal{O}_L$ over $\mathbb{F}_{p^n}$.

Remark that $T \otimes \mathbb{Z}_p = T'_p \oplus T''_p$ and F acts on T'_p by an invertible $\mathcal{O}_L$-linear endomorphism. We conclude that $F = \mathrm{Fr}_{p^n}$ acts on $T_pA(k)$ by the unique $\eta \in (\mathcal{O}_L \otimes \mathbb{Z}_p)^\times$ such that $\eta^2 - a\eta + p^n = 0$.

REMARK 6.14. $T \cong \mathcal{O}_L \oplus \mathcal{O}_L$ as $\mathcal{O}_L$-modules, hence the polarization module of the associated abelian variety with RM is just $\mathcal{D}_L^{-1}$. Replacing $\mathcal{O}_L$ by $\mathfrak{A}^{-1}$ in (6.15), we get the polarization module $\mathfrak{A}^2\mathcal{D}_L^{-1}$. It is easy to modify the construction to get any polarization module (and not just those of the form $\mathfrak{A}^2 \cdot \mathcal{D}_L^{-1}$). We will assume that henceforth. Thus we have the following

COROLLARY 6.15. *If $\mathfrak{M}$ is the coarse moduli space of abelian varieties with RM then every component of $\mathfrak{M} \otimes \overline{\mathbb{F}_p}$ contains an ordinary point.*

This follows because Corollary 6.8 gives that the components of $\mathfrak{M} \otimes \overline{\mathbb{F}_p}$, with no level structure, are in bijection with classes $\mathfrak{A} \in Cl(L)^+$. Ribet's construction gives an ordinary point on each component.

COROLLARY 6.16. *The ordinary locus is a dense open subset, hence it is irreducible in every component.*

Indeed, the Serre-Tate coordinates imply that the ordinary locus is open.

DEFINITION 6.17. A *μ_{p^n}-level structure* on an abelian scheme with RM by $\mathcal{O}_L$ $(A/S, \iota)$ is an $\mathcal{O}_L$-equivariant embedding

$$\mu_{p^n} \otimes \mathcal{D}_{L/\mathbb{Q}}^{-1} \longrightarrow A[p^n]. \tag{6.17}$$

A *μ_{p^∞}-level structure* is a compatible sequence of such embeddings.

REMARK 6.18.
1. An abelian variety with RM and a μ_{p^n}-level structure is ordinary.
2. The coarse moduli space of abelian varieties with RM together with level μ_{p^n} $(n \leq \infty)$ exists; denote it by $\mathfrak{M}(\mu_{p^n}) \longrightarrow \mathrm{Spec}(\mathbb{Z})$ for any $n \geq 0$ ($\mathfrak{M} = \mathfrak{M}(\mu_{p^0})$). If $n \geq 2$ it is a fine moduli space. Let $\mathfrak{M}(\mathbb{F}_p, \mu_{p^n})$ be the reduction modulo p of $\mathfrak{M}(\mu_{p^n})$ and let $\mathfrak{M}(\mathbb{F}_p)^{\mathrm{ord}}$ be the ordinary part of $\mathfrak{M}(\mathbb{F}_p)$. Then the fibers of

$$f : \mathfrak{M}(\mathbb{F}_p, \mu_{p^n}) \longrightarrow \mathfrak{M}(\mathbb{F}_p)^{\mathrm{ord}} \tag{6.18}$$

are principal homogeneous spaces under $(\mathcal{O}_L/p^n\mathcal{O}_L)^\times$ (for $n = \infty$, take $(\mathcal{O}_L \otimes \mathbb{Z}_p)^\times$).

THEOREM 6.19. (Ribet [**98**]) *Let $\mathfrak{A} \in Cl(L)^+$, let $\mathfrak{B}$ be the component of $\mathfrak{M}(\mathbb{F}_p)^{\mathrm{ord}}$ corresponding to $\mathfrak{A}$, and let $\mathfrak{N}$ be the preimage of $\mathfrak{B}$ under the morphism $f : \mathfrak{M}(\mathbb{F}_p, \mu_{p^n}) \longrightarrow \mathfrak{M}(\mathbb{F}_p)^{\mathrm{ord}}$. Then, $\mathfrak{N}$ is geometrically irreducible.*

PROOF. To simplify the exposition, assume that $\mathfrak{B}$ corresponds to the polarization class $\mathcal{D}_L^{-1}$. It is enough to prove the claim for $n = \infty$.

The morphism

$$f : \mathfrak{N} \longrightarrow \mathfrak{B} \tag{6.19}$$

is étale with fibers being principal homogeneous spaces under $(\mathcal{O}_L \otimes \mathbb{Z}_p)^\times$. It gives a p-adic character

$$\chi : \pi_1(\mathfrak{B} \otimes \overline{\mathbb{F}_p}) \longrightarrow (\mathcal{O}_L \otimes \mathbb{Z}_p)^\times. \tag{6.20}$$

The preimage $\mathfrak{N}$ is irreducible if and only if χ is surjective .

Now,

$$\pi_1(\mathfrak{B} \otimes \overline{\mathbb{F}_p}) = \bigcap_n \pi_1(\mathfrak{B} \otimes \mathbb{F}_{p^n}), \tag{6.21}$$

and

$$\mathcal{O}_L \otimes \mathbb{Z}_p = \varprojlim \ \mathcal{O}_L \otimes (\mathbb{Z}/p^k\mathbb{Z}). \tag{6.22}$$

Therefore, it is enough to prove that *for every k there exists an N such that for every $n \geq N$ the homomorphism $\pi_1(\mathfrak{B} \otimes \mathbb{F}_{p^n}) \xrightarrow{\chi} (\mathcal{O}_L/p^k\mathcal{O}_L)^\times$ is surjective.*

Given $\alpha \in (\mathcal{O}_L/p^k\mathcal{O}_L)^\times$, choose $a \in \mathcal{O}_L$, such that $a \equiv \alpha \mod p^k$, and choose $N > k$ such that $a^2 - 4p^n \ll 0$ for $n \geq N$. Consider the ordinary abelian variety (A, ι) over $\mathbb{F}_{p^n}$ constructed as in (6.15) from $\mathcal{O}_L, a, n$.

Let $z \in \mathfrak{B}(\mathbb{F}_{p^n})$ be the point corresponding to (A, ι) with residue field $k(z)$. Let χ_0 be the composition:

$$\pi_1(\mathrm{Spec}(k(z))) \longrightarrow \pi_1(\mathfrak{B} \otimes \mathbb{F}_{p^n}) \xrightarrow{\chi} (\mathcal{O}_L/p^k\mathcal{O}_L)^\times. \tag{6.23}$$

But

$$\pi_1(\mathrm{Spec}(k(z))) = \pi_1(\mathrm{Spec}(\mathbb{F}_{p^n})) \cong \hat{\mathbb{Z}} \tag{6.24}$$

and $< \mathrm{Fr}_{p^n} >$ is a dense subgroup of $\pi_1(\mathrm{Spec}(k(z)))$ that provides the isomorphism with $\hat{\mathbb{Z}}$. The character χ_0 is just the action of Fr_{p^n} on the fiber of $f : \mathfrak{N} \longrightarrow \mathfrak{B}$ over z. That is, it is the action of χ_0 on $\mu_{p^k} \otimes \mathcal{O}_L \hookrightarrow A[p^k]$. By duality, this is the action of Fr_{p^n} on $T_p(A/\overline{\mathbb{F}_{p^n}}) \mod p^k$. This was given by $\eta \in (\mathcal{O}_L \otimes \mathbb{Z}_p)^\times$ such that $\eta^2 - a\eta + p^n = 0$. We may read this equation mod p^k obtaining $\eta(\eta - \alpha) = 0$. We get that χ_0 acts by $\eta \equiv \alpha \mod p^k$ viewed as in $(\mathcal{O}_L/p^k\mathcal{O}_L)^\times$. □

CHAPTER 4

p-adic Elliptic Modular Forms

1. Introduction

In its simplest appearance, as defined by Serre [**102**], a p-adic modular form is a power series

$$\sum_{n=0}^{\infty} a_n q^n, \quad a_n \in \mathbb{Q}_p, \tag{1.1}$$

which is a limit, in a suitable sense, of usual modular forms with coefficients in $\mathbb{Q}$. Another approach is due to Dwork [**29**], who used p-adic analytic functions on modular curves, endowed with the action of the U-operator. Those notions were superseded and generalized by Katz, who gave a more conceptual definition of such a form as a section of a line bundle over open sets of the moduli space of elliptic curves with level structure.

In the next two sections we shall describe briefly two sources of motivation for the study of p-adic modular forms: first, p-adic L-functions, and second, deformations of Galois representations.

1.1. p-adic L-functions. Let K be a number field of degree $g = [K : \mathbb{Q}]$. Its zeta function:

$$\zeta_K(s) = \sum_{\mathfrak{a} \subset \mathcal{O}_K} \frac{1}{\mathrm{Norm}(\mathfrak{a})^s} = \prod_{\substack{\mathfrak{p} \subset \mathcal{O}_K \\ \mathfrak{p} \text{ prime}}} \frac{1}{1 - \mathrm{Norm}(\mathfrak{p})^{-s}}, \quad \Re(s) > 1, \tag{1.2}$$

can be continued meromorphically to the complex plane and satisfies the following functional equation:

$$A^s \Gamma\left(\frac{s}{2}\right)^{r_1} \Gamma(s)^{r_2} \zeta_K(s) = A^{1-s} \Gamma\left(\frac{1-s}{2}\right)^{r_1} \Gamma(1-s)^{r_2} \zeta_K(1-s), \tag{1.3}$$
$$A = 2^{-r_2} \pi^{-g/2} \sqrt{d_k}.$$

Here Γ is the complex gamma function interpolating the values $\Gamma(n+1) = n!$ for a nonnegative integer n. Note that the product in (1.2), called an *Euler product*, shows that ζ_K does not vanish for $\Re(s) > 1$.

The set of poles of the gamma function is precisely the nonpositive integers. The gamma function does not vanish on the real line. Therefore, if K is not totally real (i.e. $r_2 > 0$), then $\zeta_K(1-m) = 0$ for every integer $m > 1$; if K is totally real,

then $\zeta_K(1-m)=0$ for every odd integer $m>1$. Thus, we are interested in the numbers $\zeta_K(1-m), m\geq 2$, where m is an even integer, and K is totally real.[1]

Recall that $\zeta_K(1-m), m\geq 2$ is a rational number (See Chapter 3, Section 5.1). As a further indication that there is bound to be interesting information encoded in these values, we recall that the ubiquitous Bernoulli numbers appear as special values of the Riemann's zeta function $\zeta_{\mathbb{Q}}$:

$$\zeta_{\mathbb{Q}}(1-k)=-\frac{B_k}{k},\ \ k\geq 2. \tag{1.4}$$

Here B_k is the k-th Bernoulli number, which is zero for k odd. We include a table of some values of $\zeta_{\mathbb{Q}}$.

k	$\zeta_{\mathbb{Q}}(1-k)$		k	$\zeta_{\mathbb{Q}}(1-k)$
2	$\frac{-1}{2^2\cdot 3}$		20	$\frac{283\cdot 617}{2^3\cdot 3\cdot 5^2\cdot 11}$
4	$\frac{1}{2^3\cdot 3\cdot 5}$		22	$\frac{-131\cdot 593}{2^2\cdot 3\cdot 23}$
6	$\frac{-1}{2^2\cdot 3^2\cdot 7}$		24	$\frac{103\cdot 2294797}{2^4\cdot 3^2\cdot 5\cdot 7\cdot 13}$
8	$\frac{1}{2^4\cdot 3\cdot 5}$		26	$\frac{-657931}{2^2\cdot 3}$
10	$\frac{-1}{2^2\cdot 3\cdot 11}$		28	$\frac{9349\cdot 362903}{2^3\cdot 3\cdot 5\cdot 29}$
12	$\frac{-691}{2^3\cdot 3^2\cdot 5\cdot 7\cdot 13}$		30	$\frac{-1721\cdot 1001259881}{2^2\cdot 3^2\cdot 7\cdot 11\cdot 31}$
14	$\frac{-1}{2^2\cdot 3}$		32	$\frac{37\cdot 683\cdot 305065927}{2^6\cdot 3\cdot 5\cdot 17}$
16	$\frac{3617}{2^5\cdot 3\cdot 5\cdot 17}$		34	$\frac{-1516286975 51}{2^2\cdot 3}$
18	$\frac{-43867}{2^2\cdot 3^3\cdot 7\cdot 19}$		36	$\frac{263152715530534773 73}{2^3\cdot 3^3\cdot 5\cdot 7\cdot 13\cdot 19\cdot 37}$

Another motivation is Siegel's formula:

$$\int_{\mathrm{SL}_2(\mathcal{O}_L)\backslash\mathcal{H}^g}(-1)^n\frac{1}{(2\pi)^n}\frac{dx_1\wedge dy_1}{y_1^2}\wedge\cdots\wedge\frac{dx_n\wedge dy_n}{y_n^2}=2\zeta_L(-1), \tag{1.5}$$

for the hyperbolic volume of $\mathrm{SL}_2(\mathcal{O}_L)\backslash\mathcal{H}^g$.

Let p be a prime. Let $g=[K:\mathbb{Q}]$ and $g_p=[K\cap\mathbb{Q}(\mu_{p^\infty}):\mathbb{Q}]$, where μ_{p^∞} is the group of p-power roots of unity in $\mathbb{C}$. We have the following congruence relations:

THEOREM 1.1. (Kummer's congruences) [2]

1. *Assume* $p\geq 3, m>0$ *even.*
   a) If $g_p m\not\equiv 0 \pmod{p-1}$ *then* $\mathrm{val}_p(\zeta_K(1-m))\geq 0$.
   b) If $gm\equiv 0 \pmod{p-1}$ *then* $\mathrm{val}_p(\zeta_K(1-m))\geq -1-\mathrm{val}_p(gm)$
2. $\mathrm{val}_2(\zeta_K(1-m))\geq g-2-\mathrm{val}_2(gm)$
3. *If* $m\equiv m' \pmod{p^n(p-1)}$, *and* $gm\not\equiv 0 \pmod{p-1}$, *then*

$$(1-p^{m-1})\zeta_K(1-m)\equiv(1-p^{m'-1})\zeta_K(1-m') \pmod{p^{n+1}} \tag{1.6}$$

We shall prove certain parts of this theorem in Corollary 4.9 and in Corollary 5.3 in Chapter 5. The link to modular forms is that one can find special values, say $\zeta_K(1-m)$ or $L(1-m,\chi)$ where χ is a Dirichlet character, as leading coefficients of modular forms whose higher coefficients are *integral*. We already saw and used

[1] Those special values of the zeta function are connected to ratios of the orders of the torsion part of the K-groups of $\mathcal{O}_K$. It turns out that even when $\zeta_K(1-m)=0$, the leading coefficient of the Taylor expansion around that point still retains the K-theoretic interpretation. See [**64**]

[2] In fact this theorem can be improved, using Hilbert modular forms. See [**38**]. It can also be formulated for the prime 2.

that for the zeta function in Chapter 5.1, 6. More general constructions appear in [**23**]. See also [**115**].

One of the reasons for interest in this theorem is that it provides the p-adic interpolation of the values $\zeta_K(1-m)$ ($m \geq 2$ even) by a continuous function whose domain is $\mathbb{Z}_p$. The group $\mathbb{Z}_p$ plays an essential role in Galois theory (e.g., via the Kronecker-Weber theorem); Iwasawa has also developed a very satisfactory theory of $\mathbb{Z}_p$-extensions, using p-adic L-functions. See [**117**, Chapters 7, 13].

A more subtle link comes as follows: Let A denote the p-Sylow subgroup of $Cl(\mathbb{Q}(\zeta_p))$. Recall that p divides $|Cl(\mathbb{Q}_p)|$ iff p divides the numerator of some Bernoulli number B_k, for $k = 2, \ldots, p-3$ ([**117**, Theorem 5.16]); in that case, p is called an *irregular prime*.

EXERCISE* 1.2. Prove there are infinitely many irregular primes. (Hint: One way to prove it is to use Clausen-von Staudt theorem and Kummer congruences). It is not known whether infinitely many regular primes exist.

Let $G = \mathrm{Gal}(\mathbb{Q}(\zeta_p)/\mathbb{Q})$. The abelian group A can be considered as a $\mathbb{Z}_p[G]$-module. The group G is provided with an isomorphism to $(\mathbb{Z}/p\mathbb{Z})^\times$; a residue class n operating by $\zeta_p \mapsto \zeta_p^n$. Let ω be the Teichmüller character; it is the unique character of conductor p with values in $\mu_{p-1} \subset \mathbb{Z}_p$ such that

$$\omega(n) \equiv n \mod p, \quad n \not\equiv 0 \pmod p. \tag{1.7}$$

EXERCISE 1.3. Find the p-adic values of ω for $p = 5$.

The character ω is a generator for the character group $\hat{G}$ of G,

$$\hat{G} = \left\{\omega^i : 0 \leq i \leq p-2\right\}. \tag{1.8}$$

We decompose A in w^i-eigenspaces:

$$A = \bigoplus_{i=0}^{p-2} A_i. \tag{1.9}$$

We have $A_0 = 0$ (trivially), and $A_1 = 0$ (this needs proof). For every $i = 2, \ldots, p-2$, it is known that A_i is killed by the generalized Bernoulli number $B_{1,\omega^{-i}}$, that has the property

$$B_{1,\omega^{-i}} \equiv \frac{B_{p-i}}{p-i} \mod p. \tag{1.10}$$

We recall that for every Dirichlet character χ of conductor f, the generalized Bernoulli numbers $B_{n,\chi}$ are defined via the expansion of the following function in one variable:

$$F_\chi(t) = \sum_{a=1}^{f} \frac{\chi(a) t e^{at}}{e^{ft} - 1} = \sum_{n=0}^{\infty} B_{n,\chi} \frac{t^n}{n!}. \tag{1.11}$$

The formula (1.4) for the values $\zeta_{\mathbb{Q}}(1-m)$ can be generalized:

$$L(1-m, \chi) = -\frac{B_{m,\chi}}{m}, m \geq 1. \tag{1.12}$$

The analogue of Theorem 1.1 enables one to prove:

THEOREM 1.4. (Kubota-Leopoldt) *There exists a unique p-adic meromorphic function $L_p(s, \chi)$ (χ a Dirichlet character) with the following properties:*

1.

$$L_p(s,\chi) = \frac{a_{-1}}{s-1} + \sum_{n=0}^{\infty} a_n (s-1)^n, \quad a_n \in \mathbb{Q}(\chi) := \mathbb{Q}(\{\chi(a) : a \in \mathbb{Z}\}), \tag{1.13}$$

$$a_{-1} = \begin{cases} 1 - \frac{1}{p} & \chi = 1 \\ 0 & \chi \neq 1 \end{cases}. \tag{1.14}$$

The series converges in the disc $\{s \in \mathbb{C}_p : |s-1| < r\}, r = |p|^{\frac{1}{p-1}}|q|^{-1} > 1$ *(if* $p \neq 2$*, put* $q = p$*; otherwise, put* $q = 4$ *).*

2. *For* n *a positive integer, we have*

$$L_p(1-n,\chi) = (1 - (\chi \cdot \omega^{-n})(p) p^{n-1}) \cdot L(1-n, \chi \cdot \omega^{-n}). \tag{1.15}$$

For the proof see [**52**, Chapter 3], or [**117**, Chapter 5.2]; $\mathbb{C}_p$ is a completion of an algebraic closure of $\mathbb{Q}_p$.

1.2. Deformation of Galois representations. Let S be a finite set of primes of $\mathbb{Q}$. Let G_S be the Galois group of the maximal extension of $\mathbb{Q}$ unramified outside S, $\mathbb{F}$ a field of characteristic p, $\hat{\mathcal{C}}_{\mathbb{F}}$ the category of complete noetherian local rings $(R, \mathfrak{m}_R)$ with a given isomorphism $R/\mathfrak{m}_R \cong \mathbb{F}$.

Let $\overline{\rho} : G_S \longrightarrow \mathrm{GL}_2(\mathbb{F})$ be a representation. Given $R \in \hat{\mathcal{C}}_{\mathbb{F}}$, we say that

$$\rho : G_S \longrightarrow \mathrm{GL}_2(R) \tag{1.16}$$

is a *deformation of* $\overline{\rho}$ if the composition with the projection to $\mathbb{F}$ is $\overline{\rho}$. Two deformations ρ_1, ρ_2 over R are *equivalent* if they are conjugate by a matrix in $\mathrm{GL}_2(R) \cap (1 + M_2(\mathfrak{m}_R))$, that is, in the kernel of the map $\mathrm{GL}_2(R) \longrightarrow \mathrm{GL}_2(\mathbb{F})$ induced by the projection. One says that $\overline{\rho}$ is *ordinary* at $S_0 \subset S$, if for every inertia group I_l over a prime of S_0, the submodule $(R \times R)^{I_l}$ of vectors fixed under I_l is a free rank 1 R-module which is a direct summand of $(R \times R)^{I_l}$. A Galois representation $\overline{\rho}$ is said to be *absolutely irreducible* if there is no extension $\mathbb{F}'$ of $\mathbb{F}$ such that the representation space $\mathbb{F}'^n$ associated to $\overline{\rho} \otimes_{\mathbb{F}} \mathbb{F}'$ has a proper subspace invariant under the action of G_S.

THEOREM 1.5. (Mazur) *If* $\overline{\rho}$ *is ordinary at* S *and absolutely irreducible, then there exists a universal deformation* ρ^U *over a ring* R^U*, i.e. the deformation functor:*

$$D_{\overline{\rho}} : \hat{\mathcal{C}}_{\mathbb{F}} \longrightarrow \mathbf{Sets}, \tag{1.17}$$

is representable,

$$D_{\overline{\rho}}(R) = \mathrm{Hom}(R^U, R), \tag{1.18}$$

and $\rho^U : G_S \longrightarrow \mathrm{GL}_2(R^U)$ *is universal among all deformations of* $\overline{\rho}$*.*

Let $\rho \in S_k(\Gamma_1(n), R)$ be a newform that is an eigenvector for the Hecke operators T_l, $l \not| pN$ and the diamond operators. One can attach to f a Galois representation (Deligne, Serre, Eichler, Shimura ***):

$$\rho_f : G_S \longrightarrow \mathrm{GL}_2(R), \tag{1.19}$$

$S = \{ \text{ prime } |N\} \cup \{p, \infty\}, S_0 = \{l \in S : \overline{\rho} \text{ is ordinary at } l\}$.

Modular deformations, i.e. deformations coming from modular forms have been studied intensely in the last two decades, with some crowning achievements, such as the Shimura-Taniyama conjecture, now a theorem due mostly to Wiles, to the

effect that all elliptic curves over $\mathbb{Q}$ are modular. We recall the celebrated Serre conjecture: Let $p \geq 3$.

CONJECTURE 1.6. *Let $\rho : Gal(\overline{\mathbb{Q}}/\mathbb{Q}) \longrightarrow \mathrm{GL}_2(\mathbb{F})$ be an irreducible, odd, two-dimensional representation. There exists a normalized eigenform f of level $N(\rho)$, weight $k(\rho)$ and character $\epsilon(\rho)$ such that:*

$$\rho \sim \rho_f. \tag{1.20}$$

One important point is that the weight, level, and character are all specified. See [**103**] Sections 1-2. Excellent notes on Serre's conjecture are in preparation by Ribet and Stein [**99**].

Recall the moduli space $\mathfrak{M}(\mu_{Np^\infty})$ of elliptic curves with $\Gamma_1(N)$-structure and $\Gamma_1(\mu_{p^\infty})$-structure $(\mu_{p^n} \hookrightarrow E, \forall n)$. One let $V_{1,\infty}$ be (essentially) the space of all functions on $\mathfrak{M}(\mu_{Np^\infty})$ "defined over R". It is a space in which every space of classical modular forms $S_k(\Gamma_1(Np^\nu), R)$ injects, and their union is dense in $V_{1,\infty}$. There is a certain Hecke algebra $\mathbb{T}$ acting on $V(N, R)$ and an eigenvector f for $\mathbb{T}$ corresponding to a homomorphism $\mathbb{T} \xrightarrow{\phi_f} R \longrightarrow \mathbb{F}$; let m_f be the kernel of the composition. We call the completion of $\mathbb{T}$ at m_f the universal modular deformation $R^m(f)$.

Suppose $\overline{\rho}$ is absolutely irreducible and I_p-ordinary. Let $R(\overline{\rho})$ be the associated universal deformation ring.

CONJECTURE 1.7. *(*Mazur*) The universal deformation ring and the universal modular deformation are isomorphic:*

$$R(\overline{\rho}) \cong R^m(f), \tag{1.21}$$

i.e. all deformations of a modular residual representation are (p-adically) modular.

See [**40**] for a more complete introduction to deformations of Galois representation.

2. Congruences between Modular Forms mod p

In this section, we explore some classical congruences involving coefficients of q-expansions of modular forms. We begin our discussion trying to stay as "lowbrow" as possible, so that a story emerges out of concrete facts.

To begin with, there are striking congruences for the Fourier coefficients $\tau(n)$ of Δ:

$$\Delta(q) = q \prod_{n=1}^{\infty} (1-q^n)^{24} = \sum_{n=1}^{\infty} \tau(n) q^n. \tag{2.1}$$

We recall that Δ (discussed in Chapter 5) is the unique cusp form of weight 12 for $\mathrm{SL}_2(\mathbb{Z})$, hence an eigenform for all the Hecke operators. This already implies the relations:

(2.2)

$$\tau(mn) = \tau(m)\tau(n), \ \ (m,n) = 1; \quad \tau(p^{n+1}) = \tau(p)\tau(p^n) - p^{11}\tau(p^{n-1}), \ \ p \text{ prime}.$$

Perhaps the most elegant of the congruences satisfied by Δ, found by Ramanujan, is:

$$\tau(n) \equiv \sigma_{11}(n) \mod 691. \tag{2.3}$$

Here is a table of some values of $\tau(n)$:

n	$\tau(n)$		n	$\tau(n)$
1	1		13	-577738
2	-24		14	401856
3	252		15	1217160
4	-1472		16	987136
5	4830		17	-6905934
6	-6048		18	2727432
7	-16744		19	10661420
8	84480		20	-7109760
9	-113643		21	-4219488
10	-115920		22	-12830688
11	534612		23	18643272
12	-370944		24	21288960

EXERCISE* 2.1. What can you say on the sign of $\tau(n)$?

Ramanujan's τ function played a very important role in the development of the theory of modular forms. We just mention the consequence of Deligne's deep work [**17**], [**18**] on the Weil conjectures:

$$|\tau(n)| \leq \sigma_0(n) n^{11/2}, \tag{2.4}$$

(called "Ramanujan's conjecture") and the intriguing

CONJECTURE 2.2. (Lehmer) $\tau(n) \neq 0, \ \forall n \geq 1$.

Recall that the Eisenstein series $E_m = E_m^{\mathbb{Q}}$ for $m \geq 4$ even, are modular forms of level 1 and weight m (Chapter 5) and have the following q-expansion

$$E_m^{\mathbb{Q}} = 1 + \frac{2}{\zeta_{\mathbb{Q}}(1-m)} \sum_{n=1}^{\infty} \sigma_{m-1}(n) q^n = 1 - \frac{2m}{B_m} \sum_{n=1}^{\infty} \sigma_{m-1}(n) q^n. \tag{2.5}$$

We also define

$$E_m^* = E_m^{\mathbb{Q},*} = \frac{\zeta_{\mathbb{Q}}(1-m)}{2} E_m^{\mathbb{Q}} \tag{2.6}$$

$$= \frac{\zeta_{\mathbb{Q}}(1-m)}{2} + \sum_{n=1}^{\infty} \sigma_{m-1}(n) q^n. \tag{2.7}$$

Put

$$P = E_2 = 1 - 24 \sum \sigma_1(n) q^n \tag{2.8}$$

$$Q = E_4 = 1 + 240 \sum \sigma_3(n) q^n \tag{2.9}$$

$$R = E_6 = 1 - 504 \sum \sigma_5(n) q^n \tag{2.10}$$

Ramanujan's congruence

$$\tau(n) \equiv \sigma_{11}(n) \pmod{691} \tag{2.11}$$

is equivalent to

$$\Delta \equiv E_{12}^* \pmod{691}. \tag{2.12}$$

To see this, recall that

$$\Delta = \frac{1}{1728}(Q^3 - R^2), \tag{2.13}$$

because the right hand side is a normalized cusp form of weight 12. Granting ourselves the expression

$$E_{12}^* = \frac{1}{65520}(441Q^3 + 250R^2), \tag{2.14}$$

and noting that $441 \cdot 1728 \equiv 65520 \mod 691$, and $250 \equiv -441 \mod 691$, the congruence (2.12) follows, because the polynomials in Q, R expressing both sides are congruent mod 691. We also note that $E_{690} \equiv 1 \mod 691$. We shall see that all the congruences are of these types. That is, congruences, say modulo the prime $p = 691$, are either coming from p-integral polynomial expressions being congruent modulo 691, or from $E_{690} \equiv 1 \mod 691$.

The expression (2.14) follows from

LEMMA 2.3. *Put* $F_m = \frac{E_m^*}{(m-2)!}$. *Then, for* $m \geq 4$ *even,*

$$\frac{(m-2)(m+5)}{12}F_{m+4} = F_4F_m + F_6F_{m-2} + \cdots + F_mF_4. \tag{2.15}$$

PROOF. The Weierstrass $\wp$ function is given by

$$\wp(z,\tau) = \frac{1}{z^2} + 2\sum_{m\in 2\mathbb{N}}^{\infty}(-1)^{\frac{m+2}{2}}(2\pi)^{m+2}F_{m+2}z^m, \tag{2.16}$$

and it satisfies the differential equation

$$\wp'(z,\tau)^2 = 4\wp(z,\tau)^3 - g_2(\tau)\wp(z,\tau) - g_3(\tau), \tag{2.17}$$

where $g_2(\tau)/E_4(\tau)$ and $g_3(\tau)/E_6(\tau)$ are constant and the derivatives are taken with respect to the variable z. The second derivative is:

$$\wp''(z,\tau) = 6\wp(z,\tau)^2 - g_2(\tau)/2. \tag{2.18}$$

Compare both sides of (2.18): The coefficient of z^m on the left hand side is

$$2\cdot(-1)^{\frac{m+4}{2}}(2\pi)^{m+4}(m+1)(m+2)F_{m+4}, \tag{2.19}$$

while on the right hand side we have:

$$6\cdot 4\cdot \sum_{\substack{l+k=m,\\ l,\,k \text{ even}}}(-1)^{\frac{l+2+k+2}{2}}(2\pi)^{l+2+k+2}F_{k+2}F_{l+2} + 6\cdot 4\cdot(-1)^{\frac{m+4}{2}}(2\pi)^{m+4}F_{m+4}. \tag{2.20}$$

The Lemma follows. □

Here are some explicit expressions

$$\begin{array}{ll} E_2 = P & E_8 = Q^2 \\ E_4 = Q & E_{10} = QR \\ E_6 = R & E_{12} = \frac{1}{691}(441Q^3 + 250R^2) \\ & E_{14} = Q^2R \end{array}$$

Let $\mathcal{M} = \oplus \mathcal{M}_k$, the ring of modular forms on $\mathrm{SL}_2(\mathbb{Z})$. For every ring $B \supset \mathbb{Z}$ we let $\mathcal{M}(B) = \oplus_k \mathcal{M}_k(B)$ be the subring of modular forms whose Fourier coefficients belong to B. Let $S = \oplus_k S_k$ be the ideal of cusp forms. Recall that

$$\mathcal{M} = \mathcal{M}(\mathbb{C}) = \mathbb{C}[Q, R] \tag{2.21}$$

is a free polynomial ring in two variables and that $f \mapsto f \cdot \Delta$ yields an isomorphism

$$\mathcal{M}_k(B) \cong \mathcal{S}_{k+12}(B) \tag{2.22}$$

for every R.

LEMMA 2.4. *Let f be a modular form, and*

$$\mathcal{O}_f = \mathbb{Z}[a_0(f), a_1(f), \dots]. \tag{2.23}$$

Then f has a unique expression as an isobaric element of

$$\mathcal{O}_f[\Delta, Q] \oplus R\mathcal{O}_f[\Delta, Q]. \tag{2.24}$$

PROOF. The uniqueness is easy and is left as an exercise to the interested reader. We prove the existence by induction on the weight. For $k < 12$, it is clear. For $k \geq 12$, $k = 4a + 6b$, we obtain the cusp form $f - a_0(f)Q^a R^b = \Delta \cdot g$. We have $\mathcal{O}_g \subset \mathcal{O}_f$ (because $q^{-1}\Delta \in \mathbb{Z}[[q]]^\times$), and the result follows by induction. □

EXERCISE 2.5. Prove that $\mathcal{O}_f$ is finitely generated over $\mathbb{Z}$.

Before addressing the issue of determining all the congruences between modular forms on $\mathrm{SL}_2(\mathbb{Z})$, we discuss a certain derivation operator θ due to Ramanujan.

Let θ be the derivation

$$\theta = q\frac{d}{dq} = \frac{1}{2\pi i}\frac{d}{d\tau}. \tag{2.25}$$

Given a q-expansion $f = \sum a_n q^n$, we obtain a new q-expansion $\theta f = \sum n a_n q^n$. We shall use the following

FACT 2.6. We have

$$P(\tau) = P(\tau + 1), \quad P(-\frac{1}{\tau}) = \tau^2 P(\tau) + \frac{12\tau}{2\pi i}. \tag{2.26}$$

THEOREM 2.7. (Ramanujan) *If $f \in \mathcal{M}_k$, then*

$$\delta_k f := 12\theta f - kPf \in \mathcal{M}_{k+2}. \tag{2.27}$$

PROOF. We want to show: $(\delta_k f)(-\frac{1}{\tau}) = \tau^{k+2}\delta_k f(\tau)$.

$$\tau^{k+2}\delta_k f(\tau) = 12\tau^{k+2}\theta f(\tau) - k\left(\tau^2 \cdot P(\tau) + \frac{12\tau}{2\pi i}\right)\tau^k f(\tau) + \frac{12k}{2\pi i}\tau^{k+1} f(\tau) \tag{2.28}$$

$$= -k(Pf)(-\frac{1}{\tau}) + \frac{12k}{2\pi i}\tau^{k+1} f(\tau) + 12\tau^{k+2}\theta f(\tau) \tag{2.29}$$

On the other hand

$$(\theta f)(\tau) = \frac{1}{2\pi i}\frac{d}{d\tau}\left(\frac{1}{\tau^k} f(-\frac{1}{\tau})\right) = -\frac{k}{2\pi i}\tau^{-(k+1)} f(-\frac{1}{\tau}) + \tau^{-(k+2)}\theta f(-\frac{1}{\tau}). \tag{2.30}$$

Replacing in (2.30) τ by $-\frac{1}{\tau}$, using k even and (2.28), we get:

$$12(\theta f)(-\frac{1}{\tau}) = \frac{12k}{2\pi i}\tau^{k+1} f(\tau) + 12\tau^{k+2}(\theta f)(\tau) = \tau^{k+2}\delta_k f(\tau) + kPf(-\frac{1}{\tau}), \tag{2.31}$$

and (2.27) follows. □

DEFINITION 2.8. Define $\delta : M \longrightarrow M$, $\delta = \oplus\delta_k$. If $f \in \mathcal{M}_k, \delta f = \delta_k f = 12\theta f - kPf$.

EXERCISE 2.9. Verify the following identities:

$$\delta Q = -4R, \ \ \delta R = -6Q^2, \ \ \delta\Delta = 0, \ \ \delta P = -P^2 - Q. \tag{2.32}$$

COROLLARY 2.10. 1. *δ is the unique derivation of $\mathcal{M}$ such that*

$$\delta Q = -4R, \ \ \delta R = -6Q^2. \tag{2.33}$$

2. *$\mathbb{C}[P, Q, R]$ is stable under δ.*

PROOF. The only thing to prove is that δ is a derivation. Indeed,

$$\delta(f_1 f_2) = 12\theta(f_1 \cdot f_2) - (k_1 + k_2)Pf_1 f_2 \tag{2.34}$$

$$= (12\theta f_1 - k_1 P f_1)f_2 + (12\theta f_2 - k_2 P f_2)f_1 \tag{2.35}$$

$$= \delta(f_1)f_2 + \delta(f_2)f_1. \tag{2.36}$$

□

We now begin to study modular forms modulo p, and we look first at the q-expansion map modulo p. While in characteristic zero the q-expansion map

$$\mathcal{M}(\mathbb{C}) = \oplus_k \mathcal{M}_k(\mathbb{C}) \longrightarrow \mathbb{C}[[q]], \tag{2.37}$$

is an injective ring homomorphism, this is not true in characteristic p. As we shall see, first for $\mathrm{SL}_2(\mathbb{Z})$ (by "naive" methods) and then in greater generality, the q-expansion map is injective on each $\mathcal{M}_k$ but has a kernel on $\mathcal{M}$.

DEFINITION 2.11. Let $\mathcal{N}_k = \mathcal{M}_k(\mathbb{Z}_{(p)})$. We identify $\mathcal{N}_k$ with its image under the q-expansion map. Thus $\mathcal{N}_k$ consists of all q-expansions of modular forms of weight k on $\mathrm{SL}_2(\mathbb{Z})$ whose coefficients are p-integral. We let

$$\mathcal{N} = \oplus_k \mathcal{N}_k \hookrightarrow \mathbb{Z}_{(p)}[[q]]. \tag{2.38}$$

Let $\widetilde{\mathcal{N}}$ be the reduction of $\mathcal{N} \subset \mathbb{Z}_{(p)}[[q]]$ modulo p, so $\widetilde{\mathcal{N}} \subset \mathbb{F}_p[[q]]$.

THEOREM 2.12. *Let $p \geq 5$.*

1. *$\mathcal{N} = \mathbb{Z}_{(p)}[Q, R]$.*
2. *Let A be the polynomial such that $A(Q, R) = E_{p-1}$. Then, via the q-expansion map,* [3]

$$\widetilde{\mathcal{N}} \cong \mathbb{F}_p[Q, R]/(\widetilde{A} - 1). \tag{2.39}$$

PROOF. Lemma 2.4 gives $\mathcal{N} = \mathbb{Z}_{(p)}[Q, \Delta] \oplus R\mathbb{Z}_{(p)}[Q, \Delta]$. Note that $\Delta = \frac{1}{1728}(Q^3 - R^2) \in \mathbb{Z}_{(p)}[Q, R]$. Thus the first assertion follows.

Let

$$\mathfrak{A} = \mathrm{Ker}(\mathbb{F}_p[Q, R] \longrightarrow \widetilde{\mathcal{N}}). \tag{2.40}$$

The ideal $\mathfrak{A}$ cannot be maximal, else both $\widetilde{R} - 1$ and $\widetilde{Q} - 1$ would be algebraic over $\mathbb{F}_p$. But $\gcd(240, 504, p) = 1$, and therefore the coefficient of q in either $\widetilde{Q} - 1$ or $\widetilde{R} - 1$ is not zero, contradiction. By the Clausen - von Staudt Theorem $2\zeta(1-p)^{-1}$ is zero modulo p. Thus, E_{p-1} has q-expansion congruent to 1 modulo p. Therefore

[3] We use $\widetilde{\ }$ to denote reduction modulo p. When we write $\widetilde{Q}$ etc., we mean the reduction of the q-expansion of Q modulo p

$\widetilde{A}-1 \in \mathfrak{A}$ and since $\dim(\mathbb{F}_p[Q,R]) = 2$, we conclude that it is enough to show that $\widetilde{A}-1$ is irreducible. We shall use the following

FACT 2.13. (Igusa [**50**]) The polynomial $\widetilde{A}$ has simple roots.

We will prove this fact later, in much greater generality in fact. See Theorem 3.5.

Suppose $\widetilde{A}-1$ is reducible, then it has a factor of the form

$$\Phi(Q,R) = \Phi_n(Q,R) + \cdots + 1, \tag{2.41}$$

($n < p-1$, Φ_i = weight i part). The group $\mathbb{F}_p^\times$ acts on the graded ring $\mathbb{F}_p[Q,R]$ (where Q has weight 4 and R has weight 6): If $\zeta \in \mathbb{F}_p^*$ and g is homogenous of weight n then $[\zeta]g = \zeta^n g$. We note that $\widetilde{A}-1$ is *fixed* by this action. Therefore,

$$[\zeta]\Phi(Q,R) = \zeta^n\Phi_n(Q,R) + \cdots + 1 \tag{2.42}$$

is a *distinct* factor of $\widetilde{A}-1$, and hence $\Phi(Q,R)\cdot[\zeta]\Phi(Q,R)$ divides $(\widetilde{A}-1)$. By weight considerations, $\Phi_n(Q,R)\cdot\zeta^n\Phi_n(Q,R)$ divides $\widetilde{A}$, hence $\widetilde{A}$ has a repeated factor, contradicting Fact 2.13. □

COROLLARY 2.14. 1. *$\widetilde{\mathcal{N}}$ has a natural $\mathbb{Z}/(p-1)\mathbb{Z}$ grading,*

$$\widetilde{\mathcal{N}} = \bigoplus_{\alpha\in\mathbb{Z}/(p-1)\mathbb{Z}} \widetilde{\mathcal{N}}^\alpha, \quad \widetilde{\mathcal{N}}^\alpha = \bigcup_{k\equiv\alpha \mod p-1} \widetilde{\mathcal{N}_k}. \tag{2.43}$$

2. *Let $f \in \mathcal{N}_k$, $f' \in \mathcal{N}_{k'}$ and $f \not\equiv 0 \mod p$. Then $f \equiv f' \mod p$ implies that $k \equiv k' \mod p-1$.*

Let $A(Q,R)$, $B(Q,R)$ be the polynomials such that $A(Q,R) = E_{p-1}$ and $B(Q,R) = E_{p+1}$. Examples for $p = 5, 7, 11$ and 13 are provided in Page 101.

THEOREM 2.15. 1. $\widetilde{A}(\widetilde{Q},\widetilde{R}) = 1, \widetilde{B}(\widetilde{Q},\widetilde{R}) = \widetilde{P}$.
2. $\delta\widetilde{A} = \widetilde{B}, \delta\widetilde{B} = -Q\widetilde{A}$.
3. *The polynomials $\widetilde{A}$ and $\widetilde{B}$ are relatively prime.*
4. *The algebra $\widetilde{\mathcal{N}}$ is stable under the derivation θ.*

PROOF. First, recall that

$$E_{p+1} = 1 - \frac{2(p+1)}{B_{p+1}}\sum \sigma_p(n)q^n. \tag{2.44}$$

However, by the Kummer congruences (Theorem 1.1)

$$\frac{B_{p+1}}{p+1} \equiv \frac{B_2}{2} \mod p, \tag{2.45}$$

and obviously $\sigma_p(n) \equiv \sigma_1(n) \mod p$ for all n. Hence,

$$E_{p+1} \equiv P \pmod p. \tag{2.46}$$

We now compute the action of δ:

$$\delta\widetilde{A}(\widetilde{Q},\widetilde{R}) = 12\theta\widetilde{A}(\widetilde{Q},\widetilde{R}) - (p-1)\widetilde{P}\widetilde{A}(\widetilde{Q},\widetilde{R}) \tag{2.47}$$

$$= \widetilde{P} \tag{2.48}$$

$$= \widetilde{B}(\widetilde{Q},\widetilde{R}). \tag{2.49}$$

Therefore, $\delta A(Q,R) - B(Q,R)$ has q-expansion with coefficients in $p\mathbb{Z}_{(p)}$. Hence by Lemma 2.4, $\delta A(Q,R) - B(Q,R)$ is in $p\mathbb{Z}_{(p)}[Q,\Delta] \oplus Rp\mathbb{Z}_{(p)}[Q,\Delta] = p\mathbb{Z}_{(p)}[Q,R]$. Thus, $\delta A - B = 0 \mod p$.

We use a similar argument for B:

$$\delta\widetilde{B}(\widetilde{Q},\widetilde{R}) = 12\theta\widetilde{B}(\widetilde{Q},\widetilde{R}) - (p+1)\widetilde{P}\widetilde{B}(\widetilde{Q},\widetilde{R}) \tag{2.50}$$

$$= 12\theta\widetilde{P} - \widetilde{P}^2 \tag{2.51}$$

$$= \delta\widetilde{P} + \widetilde{P}^2 \tag{2.52}$$

$$= -\widetilde{Q} \tag{2.53}$$

$$= -\widetilde{Q}\widetilde{A}(\widetilde{Q},\widetilde{R}). \tag{2.54}$$

(We used Exercise 2.9). Thus, $\delta\widetilde{B}(\widetilde{Q},\widetilde{R}) = -\widetilde{Q}\widetilde{A}(\widetilde{Q},\widetilde{R})$, whence $\delta\widetilde{B} = -Q\widetilde{A}$.

We next prove that the polynomials $\widetilde{A}$ and $\widetilde{B}$ are relatively prime. Say

$$\widetilde{A} = \Phi(Q,R)\cdot\Psi(Q,R), \tag{2.55}$$

where necessarily $(\Phi,\Psi) = 1$. Then

$$\widetilde{B} = \delta\widetilde{A} = \delta\Phi\cdot\Psi + \delta\Psi\cdot\Phi. \tag{2.56}$$

If $\Phi|\widetilde{B}$ then $\Phi|\delta\Phi$ ($\widetilde{A}$ has no repeated factor). But $\mathrm{wt}(\delta\Phi) = \mathrm{wt}(\Phi)+2$ and we get a contradiction, because $\mathbb{F}_p[Q,R]$ has no elements of weight 2.

Finally, we see that $\widetilde{P} \in \widetilde{N}$, hence $\theta f = \frac{1}{12}(\delta f + kPf) \in \widetilde{\mathcal{N}}$. □

3. Operators and Systems of Eigenvalues

3.1. A higher brow view of modular forms in characteristic p. So far we discussed modular forms in characteristic p as the reduction modulo p of modular forms in characteristic zero, and that from two perspectives: first, as the reduction of the algebra $\mathbb{Z}_{(p)}[Q,R]$ modulo p; second, as the reduction of q-expansions $\sum a_n q^n$ of modular forms such that $a_n \in \mathbb{Z}_{(p)}$.

Those view points are clearly very restricted and artificial. Moreover, the methods use very heavily the description of modular forms on a very particular group: $\mathrm{SL}_2(\mathbb{Z})$. One would like to consider modular forms in a more intrinsic way, along the lines hinted in Chapter 1, Section 4. This point of view is due to Katz and uses in an essential way the fact that modular forms are "living" over moduli spaces.

Recall the definition of an abelian variety with RM by $\mathcal{O}_L$ (L a totally real field) with μ_n level structure: Let S be a scheme over which d_L is invertible. Then we consider triples $(A,\iota,\beta)_{/S}$, where A/S is an abelian scheme, $\iota : \mathcal{O}_L \longrightarrow \mathrm{End}_S(A)$ is an embedding of rings making the relative tangent space $\mathfrak{t}_{A/S}$ into a locally free $\mathcal{O}_L\otimes\mathcal{O}_S$ module of rank 1, and $\beta : \mu_N \otimes \mathcal{D}_L^{-1} \longrightarrow A$ is an $\mathcal{O}_L$-equivariant homomorphism. Moreover, the module of symmetric $\mathcal{O}_L$-homomorphisms $A \longrightarrow A^t$ is a projective $\mathcal{O}_L$-module of rank 1 in the étale topology.

The existence of a μ_N structure implies that the fiber of A/S, over every point of S with residue field of characteristic dividing N, is ordinary.

LEMMA 3.1. *Let $N \geq 4$. The moduli problem of abelian varieties with RM and μ_N level is rigid.*

PROOF. Let $\underline{A} = (A, \iota, \beta)_{/S}$ be an abelian variety with RM and μ_N level. We may assume S is the spectrum of an algebraically closed field. Let D be the centralizer of L in $\mathrm{End}(A) \otimes \mathbb{Q}$. It is known that D is either L, a CM field such that $D^+ = L$, or a quaternion algebra over L that is ramified everywhere at ∞. See [**8**], Lemma 6.

Let $\mathcal{O}_D = D \cap \mathrm{End}(\underline{A})$. If $\xi \in \mathcal{O}_D$ is an automorphism of A preserving the polarization, then $\xi\xi^* = 1$, where $*$ is the unique positive involution of D. Hence, ξ is of finite order. It follows that the field $L(\xi)$ is either L, or a CM field whose totally real subfield is L, and that ξ is a root of unity of order n. The case of $L(\xi) = L$ is just the case of $\xi = \pm 1$ and is easily dispensed with. We assume that $L(\xi) \neq L$. Hence, $[L(\xi) : \mathbb{Q}] = 2g$. Equivalently, $1 < \phi(n)$, $\phi(n) | 2g$ and $L \cap \mathbb{Q}(\xi) = \mathbb{Q}(\xi)^+$.

If ξ preserves a μ_N-level structure, it follows that $N^g | \deg(1 - \xi)$. Hence, n is a prime power. Say $n = \ell^r$, ℓ a prime. Then $\deg(1 - \xi) = \ell^{2g/\phi(n)}$. Since $\phi(n) > 1$, this is divisible by a g-th power if and only if $\phi(n) = 2$. On the other hand, $\phi(n) = \ell^{r-1}(\ell - 1)$. This implies $r = 1$ and $\ell = 3$, or $r = 2$ and $\ell = 2$. Both imply $N < 4$. □

It follows from the general theory of moduli spaces of abelian varieties that the moduli problem of μ_N-level is representable by a scheme $\mathfrak{M}(\mu_N)$ over $\mathbb{Z}[d_L^{-1}]$.[4] The morphism $\mathfrak{M}(\mu_N) \longrightarrow \mathbb{Z}[(Nd_L)^{-1}]$ is smooth of relative dimension g. There exists a universal object $\underline{A}^U = (A^U, \iota^U, \beta^U) \longrightarrow \mathfrak{M}(\mu_N)$.

We now restrict our attention to the case of elliptic curves. That is $L = \mathbb{Q}$. We shall later on (Chapter 5) lift the general discussion from the point we now leave it.

DEFINITION 3.2. Let $N \geq 4$. Let $\mathfrak{M}(B, \mu_N)$, B a $\mathbb{Z}[(Nd_L)^{-1}]$-algebra, denote the base change $\mathfrak{M}(\mu_N) \times_{\mathrm{Spec}(\mathbb{Z}[(Nd_L)^{-1}])} \mathrm{Spec}(B)$. It represents abelian schemes with RM and μ_N level structure over bases $S \longrightarrow \mathrm{Spec}(B)$. Let

$$\underline{A}^B = (A^B, \iota^B, \beta^B) \longrightarrow \mathfrak{M}(B, \mu_N) \tag{3.1}$$

denote its universal object (obtained by base change from $(A^U, \iota^U, \beta^U) \longrightarrow \mathfrak{M}(\mu_N)$).

A *modular form* f over B of weight k and μ_N-level is a section of the k-th tensor power of the relative cotangent space: $(\mathfrak{t}^*_{A^B/\mathfrak{M}(B,\mu_N)})^{\otimes k}$.

Using the property of the moduli space as classifying triples $(A, \iota, \beta)/S/B$, and the fact that if $(A, \iota, \beta)/S$ is parameterized by a morphism $\phi : S \longrightarrow \mathfrak{M}(B, \mu_N)$ of schemes over B then $\phi^*(\mathfrak{t}^*_{A^B/\mathfrak{M}(B,\mu_N)})^{\otimes k} = (\mathfrak{t}^*_{A/S})^{\otimes k}$ we find the following reformulation of Definition 3.2:

DEFINITION 3.3. Let $N \geq 4$. A *modular form* over B of weight k and μ_N level is a rule f associating to any triple $(A, \iota, \beta)/S/B$ an element $f((A, \iota, \beta)/S/B)$ in $(\mathfrak{t}^*_{A/S})^{\otimes k}$. The rule f is compatible with isomorphisms and commutes with base change.

Using that the sheaf $(\mathfrak{t}^*_{A/S})^{\otimes k}$ over S has the property of being locally free (in general, for RM, we shall use that condition **(R)** holds) we arrive at the following reformulation:

[4] In fact, using condition **(DP)**, we may get it over $\mathbb{Z}$.

DEFINITION 3.4. Let $N \geq 4$. A *modular form* over B (B a $\mathbb{Z}[(Nd_L)^{-1}]$-algebra) of weight k and μ_N level, is a rule f associating to any quadruple $(A, \iota, \beta, \omega)/R/B$, where R is a B-algebra, and where ω is an R-basis to $\mathfrak{t}^*_{A/R}$, an element

$$f(A, \iota, \beta, \omega) \in R \tag{3.2}$$

that depends only on the isomorphism class of $(A, \iota, \beta, \omega)/R$, commutes with base change, and satisfies

$$f(A, \iota, \beta, \alpha^{-1}\omega) = \alpha^k f(A, \iota, \beta, \omega), \quad \forall \alpha \in R^\times. \tag{3.3}$$

Unfortunately, it is quite hard to come up with a definition of a modular form using this language. This leaves the analytic methods as the most powerful methods of generating modular forms. To the rescue comes the q-expansion principle. It allows one to study modular forms defined *analytically* as *arithmetic* objects.

First, note that since modular forms are sections of one scheme over another, there is a Galois action. In particular, if B is a field then $\mathrm{Gal}(\overline{B}/B)$ acts on the modular forms defined over B.

THEOREM 3.5. *(q-expansion principle) Let $N \geq 4$ and let f be a complex modular form. Let $\sum_n a_n q^n$ be the q-expansion of f with respect to the cusp $i\infty$.*

1. *The form f^σ, $\sigma \in \mathrm{Aut}(\mathbb{C})$, has q-expansion $\sum_n a_n^\sigma q^n$; f is defined over a ring $R \subset \mathbb{C}$ if and only if $a_n \in R$ for all n.*
2. *Let $k \subset \mathbb{C}$ be a number field. Let $\mathcal{O}_k$ be its ring of integers, and let $\mathfrak{p} \lhd \mathcal{O}_k$ be a prime ideal relatively prime to Nd_L. Let f be a modular form over k. Then there exists $a \in k$ such that $aa_n \in R$ for all n. Assume that f is defined over R, then $f \pmod{p}$ (i.e., the section obtained after base change $(-) \times_R R/\mathfrak{p}$) is zero if and only if every a_n belongs to $\mathfrak{p}$.*

REMARK 3.6. Below (Chapter 5, Section 2) we introduce q-expansions in every characteristic. It would allow a sharper formulation of the q-expansion principle. Namely, that q-expansions commute with base change.

We note that the group $(\mathbb{Z}/N\mathbb{Z})^\times$ acts as automorphisms on $\mathfrak{M}(\mu_N)$. Finally, we define a modular form f over B of level 1 and weight k as a couple of modular forms g, h of levels 4 and 5 respectively, that are defined over $B[1/2]$ and $B[1/5]$ (resp.), are $(\mathbb{Z}/4\mathbb{Z})^\times$ and $(\mathbb{Z}/5\mathbb{Z})^\times$ equivariant (resp.), and agree under pull-back to $B[1/10]$.

One has to note that there could be modular forms in characteristic p that are not obtained from characteristic zero. E.g., there could be forms of weight 1 and level 1.

3.2. Operators. We discuss some operators acting on modular forms in characteristic p. We assume, for simplicity, that our modular forms are defined over a field $\mathbb{F}$.

DEFINITION 3.7. *(Hecke Operators)* For every prime ℓ, different from the characteristic of $\mathbb{F}$, we define the operator T_ℓ as follows: For every couple $(E, \omega)_{/R}$ consisting of an elliptic curve over an $\mathbb{F}$-algebra R and ω a non vanishing differential (i.e., an R-basis to $\mathfrak{t}^*_{E/R}$),

$$(T_\ell f)(E, \omega) = \frac{1}{\ell} \sum_{H < E} f(E/H, \pi_{H*}\omega). \tag{3.4}$$

The sum extends over all subgroup schemes H of E of order ℓ; for every H we denote by $\pi_H : E \longrightarrow E/H$ the natural morphism.

One verifies immediately that $T_\ell f$ is a modular form over $\mathbb{F}$ of the same weight as f, and that the Hecke operators commute with each other. Thus the free algebra $\mathbb{Z}[T_\ell : \ell \neq \mathrm{char}(\mathbb{F})]$ acts on modular forms of weight k.

If $\mathbb{F} = \mathbb{C}$ the definition of the Hecke operators takes the usual form. Let $E = \mathbb{C}/(\mathbb{Z} + \tau\mathbb{Z})$. Then the subgroups H of order ℓ corresponds to the $\ell + 1$ points in the projective space obtained from $(\ell^{-1}\mathbb{Z}/\mathbb{Z}) + (\ell^{-1}\tau\mathbb{Z}/\tau\mathbb{Z})$ that we identify with $(\mathbb{Z}/\ell\mathbb{Z})^2$. The points $\{(a : 1) : 0 \leq a \leq \ell - 1\}$ and $(1 : 0)$ account for all of them and give us the subgroups H of E that are of order ℓ. Using the obvious notation, the lattices $\mathcal{L}_H = H + (\mathbb{Z} + \tau\mathbb{Z})$ all contain the lattice $\mathbb{Z} + \tau Z$ and the natural map

$$\mathbb{C}/(\mathbb{Z} + \tau Z) \longrightarrow \mathbb{C}/\mathcal{L}_H, \tag{3.5}$$

is an isogeny with kernel H under which $dz \mapsto dz$.

Now, the point $(a : 1)$ gives us the lattice spanned by $\{1, \tau, \frac{\tau+a}{\ell}\}$, or simply by $\{1, \frac{\tau+a}{\ell}\}$. Thus, the points $\{(a : 1) : 0 \leq a \leq \ell - 1\}$ contribute to the sum in (3.4) $\frac{1}{\ell}\sum_{0\leq a\leq \ell-1} f(\frac{\tau+a}{\ell})$. The other point $(1 : 0)$ defines the lattice spanned by $\{\frac{1}{\ell}, \tau\}$. The pair $(\mathbb{C}/(\frac{1}{\ell}\mathbb{Z} + \tau\mathbb{Z}), dz)$ is isomorphic to $(\mathbb{C}/(\mathbb{Z} + \ell\tau\mathbb{Z}), \frac{1}{\ell}dz)$ and contributes to the sum $\frac{1}{\ell}f(\mathbb{C}/(\mathbb{Z} + \ell\tau\mathbb{Z}), \frac{1}{\ell}dz) = \ell^{k-1}f(\ell\tau)$. Therefore, we find that for a modular form f of weight k,

$$T_\ell f(\tau) = \ell^{k-1}f(\ell\tau) + \frac{1}{\ell}\sum_{0\leq a\leq \ell-1} f(\frac{\tau + a}{\ell}). \tag{3.6}$$

This is the classical formula for Hecke operators.

One can calculate the effect of T_ℓ on q-expansions. See [**108**] Chapter I, Proposition 10.3. If f has q-expansion $\sum_n a_n q^n$ then

$$T_\ell f(q) = \sum_n a_{\ell n} q^n + \ell^{k-1}\sum_n a_n q^{\ell n}. \tag{3.7}$$

The last formula shows, by the q-expansion principle, that the action of T_p on modular forms in characteristic p that are reduction of modular forms from characteristic zero is well-defined and its effect on q-expansions is $\sum a_n q^n \mapsto \sum a_{pn} q^n$.

One may ask if there is a more intrinsic definition of this operator. For example, if there is an analogue of Definition 3.7 for the operator T_p. First, note that a problem immediately arrises from the fact that if $\phi : E \longrightarrow E'$ is a non-separable isogeny, as the one coming from dividing by the kernel of Frobenius, then ϕ^* acts as zero on the cotangent space and one can not push-forward differentials. However, the problem is *precisely* the one created by the Frobenius morphism and can be circumvented by thinking about "Frobenius the base-change" and not "Frobenius the morphism". We therefore make the following:

Definition 3.8. Let f be a modular form of level 1 defined over a field $\mathbb{F}$ of characteristic p. We define the modular form Vf of level 1 over $\mathbb{F}$, by

$$(Vf)(E, \omega) = f(E^{(p)}, \omega^{(p)}). \tag{3.8}$$

(See Appendix A for generalities on Frobenius and Verschiebung).

For example, if E is given by a Weierstrass equation $y^2 + a_1xy + a_3y = X^3 + a_2x^2 + a_4x + a_6$ with Néron differential $\omega = dx/(2y + a_1x + a_3)$ then $E^{(p)}$ is given

by $y^2 + a_1^p xy + a_3^p y = X^3 + a_2^p x^2 + a_4^p x + a_6^p$ and $\omega^{(p)}$ is again the Néron differential $dx/(2y + a_1^p x + a_3^p)$.

Now, $Vf(E, \alpha^{-1}\omega) = f(E^{(p)}, (\alpha^{-1}\omega)^{(p)}) = f(E^{(p)}, \alpha^{-p}(\omega)^{(p)}) = \alpha^{pk} Vf(E, \omega)$. Therefore Vf is of weight pk. The effect of V on q-expansions can be computed easily once Tate objects are introduced. The q-expansion is evaluation of the modular form at a particular Tate object, and, by definition, this evaluation commutes with base change. [5] If $f(q) = \sum_n a_n q^n$ then

$$(Vf)(q) = \sum_n a_n q^{pn}. \tag{3.9}$$

To get an operator which does not raise the weight, we use the Verschiebung morphism. We make the following

DEFINITION 3.9. Let f be a modular form of level 1 defined over a field $\mathbb{F}$ of characteristic p. We define the modular form Uf of level 1 over $\mathbb{F}$, by

$$(Uf)(E, \omega) = f(E/H, \pi_{H*}\omega), \tag{3.10}$$

where H is the kernel of the Verschiebung morphism.

One sees that if f is of level 1, defined over $\mathbb{F}$, and of weight k, then so is Uf. The effect on q-expansions can be calculated to be

$$(Uf)(q) = \sum_n a_{pn} q^n. \tag{3.11}$$

Thus Uf is the reduction modulo p of the operator T_p on forms of weight greater then 1. This gives us an intrinsic characteristic p definition of T_p. Note that T_p is thus defined on all characteristic p modular forms of all weights. It is clear from the definitions that the operators U, V commute with the operators T_ℓ.

The next operator acting on modular forms in characteristic p that we mention is the operator θ discussed above. It takes modular forms of weight k to modular forms of weight $k + p + 1$ and its effect on q-expansions is

$$\sum_n a_n q^n \overset{\theta}{\longmapsto} \sum_n n a_n q^n. \tag{3.12}$$

It therefore follows that θ^{p-1} raises weight by $p^2 - 1$ and acts on q-expansions by

$$\theta^{p-1}(\sum_n a_n q^n) = \sum_{(n,p)=1} a_n q^n. \tag{3.13}$$

The following identities on q-expansions are immediate:

$$UV = 1, \quad VU = 1 - \theta^{p-1}. \tag{3.14}$$

Let us denote the operator of multiplication by the modular form $\widetilde{E_{p-1}}$, whose q-expansion is 1, by $[h]$. I.e., $[h]f = \widetilde{E_{p-1}} f$. It raises the weight by $p - 1$ and commutes with T_ℓ, U, V and θ.

[5] The base change of the Tate object $(\mathbb{G}_m/\text{periods}, dt/t)$ over $\mathbb{F}_p[[q]]$ is the "same" object but considered over $\mathbb{F}_p[[q^p]]$.

We summarize all this. Let f be a modular form over a field $\mathbb{F}$ of characteristic p, of weight k and level 1. Let

$$f(q) = \sum_n a_n q^n \tag{3.15}$$

be its q-expansion.

operator	effect on weight	effect on q-expansion
$T_\ell \ \ (\ell \neq p)$	$k \mapsto k$	$\sum_n a_{\ell n} q^n + \ell^{k-1} \sum_n a_n q^{\ell n}$
U	$k \mapsto k$	$\sum_n a_{pn} q^n$
V	$k \mapsto pk$	$\sum_n a_n q^{pn}$
θ	$k \mapsto k+p+1$	$\sum_n n a_n q^n$
$[h]$	$k \mapsto k+p-1$	$\sum_n a_n q^n$

Some relations satisfied by these operators on q-expansions are the following:

$T_\ell U = U T_\ell, \quad T_\ell V = V T_\ell;$
$\theta T_\ell = \ell T_\ell \theta;$
$UV = 1, \quad VU = 1 - \theta^{p-1};$
$U\theta = 0, \quad \theta V = 0.$

Let $\widetilde{\mathcal{N}}$ be the reduction modulo p of the q-expansions of all p-integral modular forms on $\mathrm{SL}_2(\mathbb{Z})$ as in Section 2. We note that there is a well defined action of $T_\ell, [h], \theta, V$ and U on $\widetilde{\mathcal{N}}$.

Proposition 3.10. *There is an exact sequence*

$$0 \longrightarrow \widetilde{\mathcal{N}} \xrightarrow{V} \widetilde{\mathcal{N}} \xrightarrow{\theta} \widetilde{\mathcal{N}} \xrightarrow{U} \widetilde{\mathcal{N}} \longrightarrow 0\,. \tag{3.16}$$

Proof. We already observed that $U\theta = 0$ and $\theta V = 0$. Clearly V is injective and since $UV = 1$, U is surjective. Let $f \in \mathrm{Ker}(\theta)$. Then $f = (I - \theta^{p-1})f = VUf$ and hence $f \in \mathrm{Im}(V)$. Finally, let $f \in \mathrm{Ker}(U)$. Then $f(q) = \sum_{(n,p)=1} a_n q^n$. Therefore, $f = \theta^{p-1} f$ and hence $f \in \mathrm{Im}(\theta)$. □

We note that the definition of the operators T_ℓ for $(\ell, N) = 1$, U and V extends verbatim to the case of modular forms of level μ_N, where N is prime to $\mathrm{char}(\mathbb{F})$. The operator θ extends as well: just repeat the discussion taking level μ_N into account. These operators are also defined on all modular forms in characteristic p. In addition there are the operators $< d >$ for every class $d \in (\mathbb{Z}/N\mathbb{Z})^\times$. Their action is given by

$$(< d > f)(E, \omega, \beta_N) = f(E, \omega, [d] \circ \beta_N). \tag{3.17}$$

3.3. Filtration and systems of eigenvalues. We use the notations of Section 2. We consider the reduction modulo p of modular forms of level 1, or rather their q-expansions, modulo p. Let us fix $0 \leq \alpha \leq p-1$ and view $\widetilde{\mathcal{N}}^\alpha$ as an ascending union:

$$\widetilde{\mathcal{N}}_\alpha \overset{\times E_{p-1}}{\hookrightarrow} \widetilde{\mathcal{N}}_{\alpha+p-1} \overset{\times E_{p-1}}{\hookrightarrow} \widetilde{\mathcal{N}}_{\alpha+2p-2} \overset{\times E_{p-1}}{\hookrightarrow} \cdots\,. \tag{3.18}$$

Definition 3.11. Let $f \in \widetilde{\mathcal{N}}^\alpha$. Define the *filtration* of f, $w(f)$, to be the least k such that $f \in \widetilde{\mathcal{N}}_k$.

In fact $w(f) \equiv \alpha \mod p-1$.

LEMMA 3.12. 1. *Let $f \in \mathcal{N}_k$, $f = \Phi(Q,R) \in \mathbb{Z}_{(p)}[Q,R]$, and suppose that $\widetilde{f} \neq 0$. Then*

$$w(\widetilde{f}) < k \Leftrightarrow \widetilde{A} | \widetilde{\Phi}. \tag{3.19}$$

2. *Let $f \in \widetilde{\mathcal{N}}^\alpha$. Then*

$$w(\theta f) \leq w(f) + p + 1, \tag{3.20}$$

with equality iff $w(f) \not\equiv 0 \mod p$.

3. *Let $f \in \widetilde{\mathcal{N}}^\alpha$. Then*

$$w(Vf) = pw(f). \tag{3.21}$$

4. *Let $f \in \widetilde{\mathcal{N}}^\alpha$, then*

$$w(Uf) \leq \frac{1}{p}(w(f) + p^2 - 1), \tag{3.22}$$

with equality iff $w(f) \equiv 1 \mod p$.

PROOF. 1. Say $\widetilde{\Phi} = \widetilde{A}\widetilde{\Psi}$. Let $\Psi \in \mathbb{Z}_{(p)}[Q,R]$ be any isobaric lift of $\widetilde{\Psi}$ and $g = \Psi(Q,R)$. Then $\widetilde{g(q)} = \widetilde{f(q)}$, and $\mathrm{wt}(\widetilde{\Psi}) < \mathrm{wt}(\widetilde{\Phi})$, hence $w(\widetilde{f}) < k$.

Conversely, if $w(\widetilde{f}) < k$ let g be a modular form of weight $w(\widetilde{f})$ such that $\widetilde{g(q)} = \widetilde{f(q)}$. Write $g = \Psi(Q,R)$. Then $\widetilde{g(q)} = \widetilde{f(q)}$ implies $(\widetilde{A}-1)|(\widetilde{\Phi} - \widetilde{\Psi})$. Since $\mathrm{wt}(\widetilde{\Psi}) < \mathrm{wt}(\widetilde{\Phi})$, we get $\widetilde{A}|\widetilde{\Phi}$.

2. Let $g \in \mathcal{N}_{w(f)}$ such that $\widetilde{g} = f$. Say $g = \Phi(Q,R)$. Let $k = \mathrm{wt}(f)$. Recall that

$$12\theta\widetilde{f} = 12\theta\widetilde{g} = \omega(f)\widetilde{E_{p+1}g} + \widetilde{E_{p-1}\delta g} \in \widetilde{\mathcal{N}_{\omega(f)+p+1}}. \tag{3.23}$$

Then

$$\begin{aligned} w(\theta f) < w(f) + p + 1 &\iff \widetilde{A} | \widetilde{\omega(f)B\Phi + A\delta\Phi} \\ &\iff \widetilde{A} | \widetilde{\omega(f)B\Phi} \\ &\underset{(\widetilde{A},\widetilde{B})=1}{\iff} \widetilde{A} | \widetilde{\omega(f)}\widetilde{\Phi} \\ &\iff \text{either } w(f) \equiv 0 \mod p \text{ or } \widetilde{A}|\widetilde{\Phi}, \end{aligned} \tag{3.24}$$

but the latter is impossible, because of Part 1.

3. We write $f = \widetilde{\Phi}(\widetilde{Q},\widetilde{R})$ with $\deg(\Phi) = w(f)$, and by Part 1 $\widetilde{A} \not| \widetilde{\Phi}$. Clearly, $Vf = \widetilde{\Phi^p}(\widetilde{Q},\widetilde{R})$ and thus $w(Vf) \leq pw(f)$. But again, Part 1 implies that $w(Vf) < pw(f) \iff \widetilde{A}|\widetilde{\Phi^p}$. This cannot happen, because $\widetilde{A}$ has *simple* factors.

4. Part 2 implies that $w(\theta^{p-1}f) \leq w(f) + p^2 - 1$. The inequality is strict iff there exists a j, $0 \leq j < p-1$ such that $p|(w(f) + j(p+1))$; equivalently, $w(f) \not\equiv 1 \mod p$.

Now, $VUf = f - \theta^{p-1}f$, and using Part 3 we get

$$pw(Uf) = w(VUf) = w(f - \theta^{p-1}f) \leq \max\left\{w(f), w(\theta^{p-1}f)\right\}. \tag{3.25}$$

This gives the inequality $w(Uf) \leq \frac{1}{p}(w(f) + p^2 - 1)$. The equality occurs iff $w(f) \equiv 1 \mod p$.

□

COROLLARY 3.13. 1. *If* $w(f) \equiv 1 \mod p$, $Uf \neq 0$.
2. *If* $w(f) > p+1$, *then* $w(Uf) < w(f)$.

DEFINITION 3.14. Let $W_k = \widetilde{\mathcal{N}_k} \otimes \overline{\mathbb{F}_p} / \widetilde{\mathcal{N}_{k-p+1}} \otimes \overline{\mathbb{F}_p}$.

We note that Corollary 3.13 gives the following: If $k > p+1$, then U annihilates W_k.

We now focus our attention on the structure of $\widetilde{N}$ as a Hecke algebra. The first question is about how many different systems of eigenvalues exist modulo p? To that end we first make a

DEFINITION 3.15. A set $\{\lambda_\ell : \lambda_\ell \in \overline{\mathbb{F}_p}, \ell \neq p \text{ prime}\}$ is called *a system of eigenvalues* if there exists $0 \neq f \in \mathcal{N} \otimes \overline{\mathbb{F}_p}$ such that $T_\ell f = \lambda_\ell f$ for all $\ell \neq p$.

Even though we have infinitely many such systems in characteristic zero, in characteristic p, the situation is different:

THEOREM 3.16. (Jochnowitz)[**55**] *There exist only finitely many systems of eigenvalues.*[6]

PROOF. It is enough to prove the finiteness of systems for all the W_k's. Given $1 \leq a \leq p-1$, we define the twisted Hecke module

$$W_k[a] := W_k \otimes_{\overline{\mathbb{F}}} \overline{\mathbb{F}}, \tag{3.26}$$

where T_ℓ acts on $f \otimes r$ by $T_\ell(f) \otimes \ell^a r$.

Note that if $\{\lambda_\ell\}$ is a system in W_k then $\{\lambda_\ell \ell^a\}$ is a system in $W_k[a]$. Therefore, it is enough prove the following
Claim: If $j > 2p$ then $W_j \cong W_m[a]$ as Hecke modules for some $1 \leq a \leq p-1$ and $m < j$.

Case 1. $j \not\equiv 1 \mod p$.
Lemma 3.12 states that $w(\theta f) \leq w(f) + p + 1$ with equality if $w(f) \not\equiv 0 \mod p$. Thus, $\theta : W_k \hookrightarrow W_{k+p+1}$ if $w(f) \not\equiv 0 \mod p$. If $k \geq p+1$, this is an isomorphism. This follows from calculating the dimensions of the two vector spaces. A verification we leave to the reader.

Let us apply these observations for $k = j - p - 1$ noting that $\theta \circ T_\ell = \ell T_\ell \circ \theta$. Since $k \geq p+1$ and $k \not\equiv 0 \mod p$, we get $W_{j-p-1}[1] \cong W_j$.
Case 2. $j \equiv 1 \mod p$.

In this case, since $j > 2p > p+1$, by Corollary 3.13 and Lemma 3.12, U induces a homomorphism:

$$U : W_j \hookrightarrow W_{\frac{j-1}{p}+p}. \tag{3.27}$$

Composing with $V : W_{\frac{j-1}{p}+p} \hookrightarrow W_{j+p^2-1}$, we get

$$V \circ U : W_j \hookrightarrow W_{j+(p+1)(p-1)}. \tag{3.28}$$

This map must be an isomorphism by the same dimension count. Therefore, U is an isomorphism. □

[6]This theorem was much superseded by the results of Ash-Stevens. See [**1**], [**2**].

3.4. Congruences mod p^m. Let $p \geq 5, f \in \mathbb{Q}[Q,R]$. Let $f(q) = \sum a_n q^n$ be the q-expansion of f. By assumption, $a_n \in \mathbb{Q}$. Define

$$\mathrm{val}_p(f) := \inf\{\mathrm{val}_p(a_n) : n = 0, 1, 2, \dots\}. \tag{3.29}$$

It follows from the q-expansion principle 3.5 that $\mathrm{val}_p(f) > -\infty$.

THEOREM 3.17. (Serre) *Let $m \geq 1$. Let $f, f' \in \mathbb{Q}[Q,R]$ of weight k, k', respectively. Suppose $f \neq 0$. If $\mathrm{val}_p(f - f') \geq \mathrm{val}_p(f) + m$, then*

$$k' \equiv k \mod (p-1)p^{m-1}. \tag{3.30}$$

PROOF. By multiplying by $p^{-\mathrm{val}_p(f)}$, we may assume that $\mathrm{val}_p(f) = 0$ and hence we are given that $\mathrm{val}_p(f - f') \geq m$. Namely, both f and f' are p-integral and

$$f \equiv f' \mod p^m, \quad m \geq 1. \tag{3.31}$$

In particular, $f \equiv f' \pmod{p}$ and by Corollary 2.14

$$k' \equiv k \mod p-1. \tag{3.32}$$

Hence, the theorem is true for $m = 1$. Assume now that $m \geq 2$. We shall use the following

EXERCISE 3.18.

$$E_r \equiv 1 \mod p^m \iff r \equiv 0 \mod (p-1)p^{m-1}. \tag{3.33}$$

Let $h = k' - k$. Replacing f' by $f' E_{(p-1)p^n}(n \gg 0)$ we may assume that $h \geq 4$. We know that $h \equiv 0 \mod p-1$. Let $r = \mathrm{val}_p(h) + 1$. We want to show that $r \geq m$.

Suppose that $r < m$. We claim that $p^{-r}(fE_h - f')$ is p-integral and

$$p^{-r}(fE_h - f') \equiv p^{-r} f(E_h - 1) \mod p. \tag{3.34}$$

We have

$$fE_h - f' = f - f' + f(E_h - 1). \tag{3.35}$$

Now, $f - f' \equiv 0 \pmod{p^m}$ and $E_h - 1 \equiv 0 \pmod{p^r}$. Thus $p^{-r}(fE_h - f')$ is p-integral, and $r < m$ implies (3.34).

We have:

$$p^{-r}(E_h - 1) = \lambda\phi, \tag{3.36}$$

where $\lambda \in \mathbb{Z}_p^\times$ and $\phi = \sum_{n=1}^{\infty} \sigma_{h-1}(n) q^n$. Let

$$g := \lambda^{-1} p^{-r}(fE_h - f'). \tag{3.37}$$

Then (3.34) implies that

$$g = f\phi \mod p. \tag{3.38}$$

Put:

$$\widetilde{\phi} = \widetilde{g}/\widetilde{f}. \tag{3.39}$$

Note that $\widetilde{\phi}$ is well-defined since $\widetilde{f}$ is not zero. Since $\mathrm{wt}(g) = k'$ is congruent to $\mathrm{wt}(f) = k$ modulo $p-1$, $\widetilde{\phi}$ belongs to the quotient field of $\widetilde{\mathcal{N}}^0$. We shall use the following fact.

FACT 3.19. (See Corollary 5.4) $\widetilde{\mathcal{N}}^0$ is a Dedekind domain.

We claim that $\widetilde{\phi}$ is integral over $\widetilde{\mathcal{N}}^0$. To see this, first note that

$$\widetilde{\phi} - \widetilde{\phi}^p \equiv \sum_{(p,n)=1} \sigma_{h-1}(n) q^n \mod p. \tag{3.40}$$

Define, for $r \in \mathbb{Z}$, $(n,p)=1$,

$$\widetilde{\sigma}_r(n) = \sum_{d|n} d^r \mod p. \tag{3.41}$$

Then

$$\widetilde{\sigma}_{-1}(n) = \widetilde{\sigma}_1(n)/n, \quad \widetilde{\sigma}_r(n) = \widetilde{\sigma}_{pr}(n) = \widetilde{\sigma}_{r+p-1}(n). \tag{3.42}$$

Consider the following identity:

$$\begin{aligned}
\theta^{h-1}\Big(\sum_{n=1}^{\infty} \sigma_1(n) q^n\Big) &\equiv \sum_{n=1}^{\infty} n^{h-1} \widetilde{\sigma}_1(n) q^n \\
&\equiv \sum_{(n,p)=1} \frac{\widetilde{\sigma}_1(n)}{n} q^n \quad (h \equiv 0 \pmod{p-1}) \\
&= \sum_{(n,p)=1} \widetilde{\sigma}_{-1}(n) q^n \\
&\equiv \sum_{(n,p)=1} \widetilde{\sigma}_{h-1}(n) q^n \quad (h \equiv 0 \pmod{p-1}) \\
&= \widetilde{\phi} - \widetilde{\phi}^p.
\end{aligned} \tag{3.43}$$

Therefore,

$$\widetilde{\phi} - \widetilde{\phi^p} = -\frac{1}{24}\theta^{h-1}(\widetilde{P}) = -\frac{1}{24}\theta^{h-1}(\widetilde{E_{p+1}}) = -\frac{1}{24}\theta^{p-2}(\widetilde{E_{p+1}}), \tag{3.44}$$

(θ^{p-1} is idempotent). The last expression shows that $\widetilde{\phi} - \widetilde{\phi^p}$ belongs to

$$\widetilde{\mathcal{N}_{p+1+(p-2)(p+1)}} = \widetilde{\mathcal{N}_{p^2-1}} \subset \widetilde{\mathcal{N}}^0. \tag{3.45}$$

This proves the claim of integrality, hence that $\widetilde{\phi} - \widetilde{\phi}^p$ belongs to $\widetilde{\mathcal{N}}^0$, plus the fact that

$$\widetilde{\phi} - \widetilde{\phi}^p = -\frac{1}{24}\theta^{p-2}(\widetilde{E_{p+1}}). \tag{3.46}$$

Now, the filtration of $\widetilde{\phi} - \widetilde{\phi^p}$ is $\max\left\{\omega(\widetilde{\phi}), \omega(\widetilde{\phi^p})\right\} = \max\left\{\omega(\widetilde{\phi}), p \cdot \omega(\widetilde{\phi})\right\} = p \cdot \omega(\widetilde{\phi})$. But Lemma 3.12 says that $\omega(\theta^{p-2}\widetilde{E_{p+1}}) = p^2 - 1$, and this is a contradiction. □

4. Serre's p-adic Modular Forms and p-adic Zeta Functions

Throughout this section $p \geq 5$.

Let m be a positive integer. Put

$$X_m = \mathbb{Z}/(p^{m-1}(p-1))\mathbb{Z} = \mathbb{Z}/p^{m-1}\mathbb{Z} \times \mathbb{Z}/(p-1)\mathbb{Z}, \tag{4.1}$$

and

$$X = \varprojlim X_m = \mathbb{Z}_p \times \mathbb{Z}/(p-1)\mathbb{Z}. \tag{4.2}$$

Since the maps $\mathbb{Z} \longrightarrow X_m$ are surjective, we get that the injection $\mathbb{Z} \hookrightarrow X$ has a dense image. Define for $v = v_1 v_2 \in \mathbb{Z}_p^\times = U_1 \times \mu_{p-1}$ and $k = (k_1, k_2) \in \mathbb{Z}_p \times \mathbb{Z}/(p-1)\mathbb{Z}$

$$v^k := v_1^{k_1} v_2^{k_2} \in \mathbb{Z}_p^\times. \tag{4.3}$$

In this way

$$X = \mathrm{End}_{\mathrm{cont}}(\mathbb{Z}_p^\times). \tag{4.4}$$

DEFINITION 4.1. We say $k = k_1 k_2$ in X is *even* if $k \in 2X$. Equivalently, k_2 is even or $(-1)^k = 1$.

DEFINITION 4.2. A *p-adic modular form* (à la Serre) is a formal power series $f = \sum a_n q^n$, where:

1. $a_n \in \mathbb{Q}_p$.
2. There exists a sequence $(f_i)_{i\in\mathbb{N}}$ of modular forms of level 1 and weight k_i with rational Fourier coefficients, such that

$$\mathrm{val}_p(f - f_i) \xrightarrow{i \longrightarrow \infty} \infty. \tag{4.5}$$

THEOREM 4.3. (Serre) *The weights k_i have a limit in X. This limit depends only on f and not on the particular sequence $(f_i)_{i\in\mathbb{N}}$. It is called the weight of f.*

PROOF. For $i \gg 0$, $\mathrm{val}_p(f_i) = \mathrm{val}_p(f)$ (uniform convergence), and given any m for $j \gg i \gg 0$, we have

$$\mathrm{val}_p(f_i - f_j) \geq \mathrm{val}_p(f) + m = \mathrm{val}_p(f_i) + m. \tag{4.6}$$

Since by Theorem 3.17 $k_i \equiv k_j \mod (p-1)p^{m-1}$, there exists a limit to the k_i's.

If $f_1', f_2', \ldots$ is another sequence converging to f, consider the sequence

$$f_1, f_1', f_2, f_2', \ldots! \tag{4.7}$$

□

The weight of f is always even simply because it is a limit of even weights. Note that under val_p, the space of p-adic modular forms is a p-adic Banach space. That is, it is a vector space over $\mathbb{Q}_p$ with a norm $p^{-\mathrm{val}_p(f)}$ and the space is complete with respect to this norm. That means that if f_i are p-adic modular forms forming a Cauchy sequence with respect to the norm, then the limit of the f_i's exists.

EXERCISE 4.4. Prove that $\frac{1}{E_{p-1}} = \lim_m E_{p-1}^{p^m-1}$.

THEOREM 4.5. *Let m be a positive integer. Let f, f' be two p-adic modular forms of weight $k, k' \in X$ respectively and assume $f \neq 0$. Then $\mathrm{val}_p(f - f') \geq \mathrm{val}_p(f) + m$ implies that $k = k'$ in $X_m = \mathbb{Z}/p^{m-1}\mathbb{Z} \times \mathbb{Z}/(p-1)\mathbb{Z}$.*

EXERCISE 4.6. Prove the Theorem 4.5.

The following corollary is one of the most amazing results of the theory of p-adic modular forms. It is an integral analog of Siegel's Theorem on the rationality of values of zeta function (Chapter 2, Remark 5.5). Recall that the idea behind the proof of that theorem was that if all the higher coefficients of a modular form are rational so is the leading coefficient. Thus there is a "rational" influence of the higher coefficients on the first coefficient. The following corollary says that this influence can be refined to an integral influence: If all the higher coefficients are p-adic integers then the leading coefficient has a valuation bounded from below in terms of the weight alone!

COROLLARY 4.7. *Let $f = a_0 + a_1 q + a_2 q^2 + \cdots$ be a p-adic modular form of weight $k \neq 0$. Let $m \geq 0$ be such that $k \neq 0$ in X_{m+1}. Then*

$$\mathrm{val}_p(a_0) + m \geq \inf_{n \geq 1} \{\mathrm{val}_p(a_n)\}. \tag{4.8}$$

PROOF. If $a_0 = 0$ then $\mathrm{val}_p(a_0) = \infty$ and the assertion holds. Else $a_0 \neq 0$ and we let $f' := a_0$ be a modular form of weight 0. We have $\mathrm{val}_p(f - f') = \inf_{n \geq 1} \{\mathrm{val}_p(a_n)\}$. If, on the contrary, $\inf_{n \geq 1} \{\mathrm{val}_p(a_n)\} > \mathrm{val}_p(a_0) + m$, i.e., if $\mathrm{val}_p(f - f') \geq \mathrm{val}_p(f') + (m+1)$, then Theorem 4.5 implies that $k \equiv 0$ in X_{m+1}, a contradiction. □

COROLLARY 4.8. *Let k be such that $(p-1) \nmid k$. If a_i are integral for every $i \geq 1$ then a_0 is also integral.*

PROOF. Take $m = 0$ in Corollary 4.7. □

COROLLARY 4.9. *Let L be a totally real field of degree g. For p such that $kg \not\equiv 0 \mod p-1$, we have that $\zeta_L(1-k)$ is p-integral. More generally,*

$$\mathrm{val}_p(\zeta_L(1-k)) \geq -1 - \mathrm{val}_p(kg). \tag{4.9}$$

PROOF. Recall the modular form of level 1 and weight kg defined in Chapter 2, Section 6:

$$\Phi^*(2^{-g}\zeta_L(1-k)E_{k,\mathcal{D}_L^{-1}}) = \frac{\zeta_L(1-k)}{2^g} + \sum_{n=1}^{\infty}\Big(\sum_{\nu \in \mathcal{D}_L^{-1+}} \sigma_{k-1}((\nu)\mathcal{D}_L)\Big)q^n, \tag{4.10}$$

where $\Phi : \mathcal{H} \longrightarrow \mathcal{H}^g$ is the diagonal map.

The leading coefficient of this modular form is $\frac{\zeta_L(1-k)}{2^g}$, and the higher coefficients are p-integral (in fact, integers).

If $kg \not\equiv 0 \mod p-1$, then Corollary 4.8 says that the leading coefficient is also p-integral.

To get the general inequality, we may assume $\widetilde{k} = kg$ is congruent to 0 modulo $p-1$. Let $m = \mathrm{val}_p(\widetilde{k}) + 1$; then $\widetilde{k} \neq 0$ in X_{m+1}, hence

$$\mathrm{val}_p(\zeta_L(1-k)) + m \geq \inf\{\mathrm{val}_p(a_n)\} \geq 0. \tag{4.11}$$

□

COROLLARY 4.10. *Let*

$$f^{(i)} = \sum_{n=0}^{\infty} a_n^{(i)} q^n, \quad i = 1, 2, \ldots \tag{4.12}$$

be a sequence of p-adic modular forms of weight $k^{(i)}$. Suppose:

1. *For $n \geq 1$, $a_n^{(i)} \longrightarrow a_n \in \mathbb{Q}_p$ uniformly in n. (That is, there exists a power series limit in the val_p-topology, $\lim_i(\sum_{n=1}^{\infty} a_n^{(i)} q^n)$).*
2. *$k^{(i)} \longrightarrow k \in X$, and $k \neq 0$.*

Then $a_0^{(i)} \longrightarrow a_0 \in \mathbb{Q}_p$ and

$$f = \sum_{n=0}^{\infty} a_n q^n \tag{4.13}$$

is a p-adic modular form of weight k.

PROOF. By deleting some $f^{(i)}$'s, we may assume that there exists an $m \geq 0$ such that all the $k^{(i)}$'s have non zero image in X_{m+1}. Now, uniform convergence implies there is a $t \in \mathbb{Z}$ such that

$$\mathrm{val}_p(a_n^{(i)}) \geq t, \quad \forall n, \forall i \geq 1. \tag{4.14}$$

By Corollary 4.7 $\mathrm{val}_p(a_0^{(i)}) \geq t - m$. Hence, there exists a subsequence i_j such that $a_0^{(i_j)} \xrightarrow{j \longrightarrow \infty} a_0 \in \mathbb{Q}_p$. Then clearly,

$$f = \lim f^{(i_j)} = a_0 + a_1 q + a_2 q^2 + \dots \ . \tag{4.15}$$

is a p-adic modular form of weight k.

Remark that in a compact metric space a sequence converges if and only if every converging subsequence converges to the same limit. If i_j is another converging subsequence $a_0^{(i_j)} \longrightarrow a_0$, then $f = a_0' + \sum_{n=1}^{\infty} a_n q^n$ is a p-adic modular form of weight k, hence

$$f - f = a_0 - a_0' \tag{4.16}$$

of weight k, but also of weight zero! Since $k \neq 0$ we must have $a_0 - a_0' = 0$. □

EXAMPLE 4.11. **p-adic Eisenstein series.** Let $k \in X$, and define

$$\sigma_{k-1}^*(n) = \sum_{\substack{d|n \\ (d,\,p)\,=\,1}} d^{k-1}, \quad n \geq 1, n \in \mathbb{Z}. \tag{4.17}$$

Assume k is even, choose $k_i \geq 4$ such that $k_i \longrightarrow k$ in X, and $k_i \longrightarrow \infty$ in $\mathbb{R}$. Then $\sigma_{k-1}(n) \longrightarrow \sigma_{k-1}^*(n)$ in $\mathbb{Z}_p$ uniformly in n. Therefore, by Corollary 4.10, the modular forms G_{k_i} converge p-adically to a p-adic modular form G_k^*. Moreover,

$$G_k^* = \frac{1}{2}\zeta_p^*(1-k) + \sum_{n=1}^{\infty} \sigma_{k-1}^*(n) q^n \tag{4.18}$$

is a p-adic modular form of weight k, where we define

$$\zeta_p^*(1-k) = \lim_i \ \zeta_{\mathbb{Q}}(1-k_i). \tag{4.19}$$

(The existence of the limit is a consequence of Corollary 4.10!) One calls G_k^* the *p-adic Eisenstein series of weight k*. Note that even if k is an integer $G_k^* \neq G_k$.

REMARK 4.12. It is known that the Eisenstein series E_2 of weight 2 is not a classical modular form. But for any prime p, the q-expansion of E_2 indeed arises from a p-adic modular form.

THEOREM 4.13. (Serre) *Put $h = (s, n) \in X$ (odd). Then*

$$\zeta_p^*(1 - (s,n)) = L_p(s, \omega^{1-n}), \tag{4.20}$$

where ω is the Teichmüller character.

Recall that $L_p(s, \chi)$ interpolates the special values of classical L function $L(s, \chi)$. Thus Serre's construction creates the "correct" p-adic zeta functions: they interpolate p-adically special values of classical L functions. One can generalize Serre's construction to get a p-adic zeta function for L totally real (see Serre's paper in [**101**]).

5. A Geometric Approach to Congruences

The methods of the previous sections rely very much, so it seems, on the specific structure of the ring of modular forms on $\mathrm{SL}_2(\mathbb{Z})$. To obtain a more general theory that works for level structures as well as for Hilbert modular forms, we reconsider the question of congruences. This time from a purely characteristic p approach. Thus our original object of study are modular forms in characteristic p and not the reduction of modular forms from characteristic zero. Already for elliptic modular forms there is a difference, but for Hilbert modular forms the difference becomes much more dramatic, with important characteristic p modular forms that cannot be lifted to characteristic zero.

We shall only treat the case of elliptic curves here, allowing ourselves a simpler picture then in the general Hilbert modular case that would be discussed in Chapter 5.

Let $N \geq 4$ be an integer. Let $\mathfrak{M} = \mathfrak{M}(\mathbb{F}_p, \mu_N)$ be the fine moduli space parameterizing elliptic curves E/R over $\mathbb{F}_p$-algebras R, endowed with a μ_N level structure:

$$\beta_N : \mu_{N/R} \hookrightarrow E/R. \tag{5.1}$$

We may compactify $\mathfrak{M}$ to a scheme $\mathfrak{M}^*$ by adding finitely many cusps. In fact, $\mathfrak{M}^*$ represents a similar modular problem involving generalized elliptic curves as in [**22**]. The scheme $\mathfrak{M}^*$ is integral, geometrically irreducible, regular proper scheme of dimension 1.

Let $\mathfrak{M}(\mu_p) = \mathfrak{M}(\mathbb{F}_p, \mu_{Np})$ be the fine moduli space representing elliptic curves E/R over $\mathbb{F}_p$-algebras R, endowed with a μ_{Np} level structure:

$$\beta_N \times \beta_p : \mu_{N/R} \times \mu_{p/R} \hookrightarrow E/R. \tag{5.2}$$

Let $\mathfrak{M}^*(\mu_p)$ be the scheme obtained by adding the cusps. The scheme $\mathfrak{M}^*(\mu_p)$ is not proper. In fact, if we let $\mathfrak{M}^{\mathrm{ord}}$ denote the open subscheme of $\mathfrak{M}^*$ obtained by deleting the points corresponding to the supersingular elliptic curves then we have a surjective étale Galois morphism

$$\mathfrak{M}^*(\mu_p) \longrightarrow \mathfrak{M}^{\mathrm{ord}}, \tag{5.3}$$

with Galois group $(\mathbb{Z}/p\mathbb{Z})^\times$.

5.1. The Hasse invariant. Let $(E, \beta_N, \omega)_{/R}$ be an elliptic curve over an $\mathbb{F}_p$-algebra R, with μ_N-level β_N and a non-vanishing differential ω. We define a modular form H, called the Hasse invariant, as follows. We consider the exact sequence

$$0 \longrightarrow H^0(E, \Omega^1_{E/R}) \longrightarrow H^1_{dR}(E) \longrightarrow H^1(E, \mathcal{O}_E) \longrightarrow 0\,. \tag{5.4}$$

There is a perfect pairing

$$H^0(E, \Omega^1_{E/R}) \times H^1(E, \mathcal{O}_E) \xrightarrow{\cup} H^1(E, \Omega^1_{E/R}) \cong H^0(E, \mathcal{O}_E)^* = R \tag{5.5}$$

(using Serre's duality) and we let $\eta \in H^1(E, \mathcal{O}_E)$ be the element dual to ω under this pairing.

The sheaf $\mathcal{O}_E$ is a sheaf of rings of characteristic p and therefore the Frobenius map $f \mapsto f^p$ induces a homomorphism of sheaves of abelian groups $\mathcal{O}_E \longrightarrow \mathcal{O}_E$

and hence a map (called Frobenius) on cohomology

$$\mathrm{Fr} : H^1(E, \mathcal{O}_E) \longrightarrow H^1(E, \mathcal{O}_E), \tag{5.6}$$

which is Frobenius linear: $\mathrm{Fr}(\alpha v) = \alpha^p \mathrm{Fr}(v)$ for $\alpha \in R$ and $v \in H^1(E, \mathcal{O}_E)$. We let $H(E, \beta_N, \omega)$ be the unique element of R such that

$$\mathrm{Fr}(\eta) = H(E, \beta_N, \omega) \cdot \eta. \tag{5.7}$$

One easily verifies that the definition commutes with base change and depends only on the isomorphism class of (E, ω). Moreover, if $\alpha \in R^\times$ then the element dual to $\alpha^{-1}\omega$ is $\alpha\eta$ and $\mathrm{Fr}(\alpha\eta) = \alpha^p \mathrm{Fr}(\eta) = \alpha^{p-1} H(E, \omega)\alpha\eta$. That is,

$$H(E, \beta_N, \alpha^{-1}\omega) = \alpha^{p-1} H(E, \beta_N, \omega). \tag{5.8}$$

Thus, we proved the first half of the following proposition. The second part follows from the more general discussion of Tate objects given in Chapter 5, Section 2.

PROPOSITION 5.1. *The Hasse invariant H is a modular form over $\mathbb{F}_p$ of level one and weight $p-1$. Its q-expansion at every cusp is* 1.

5.2. The kernel of the q-expansion. Let $\mathcal{M}(\mathbb{F}_p, k, \mu_N)$ ($N \geq 4$ prime to p) be the vector space of modular forms defined over $\mathbb{F}_p$, of weight k and level μ_N.

THEOREM 5.2. (Serre – Swinnerton-Dyer) *The kernel of the q-expansion map*

$$\bigoplus_{k \in \mathbb{Z}} \mathcal{M}(\mathbb{F}_p, k, \mu_N) \longrightarrow \mathbb{F}_p[[q]], \tag{5.9}$$

is generated by $H - 1$.

PROOF. Let R be the ring of regular functions on the curve $\mathfrak{M}^*(\mu_p)$. Let $\mathbb{G}_m/q(\mathbb{Z})$ be a Tate object represented by a cusp **Tate** of $\mathfrak{M}^*(\mu_p)$.

PROPOSITION 5.3. *There is a surjective ring homomorphism*

$$r : \bigoplus_{k \in \mathbb{Z}} \mathcal{M}(\mathbb{F}_p, k, \mu_N) \twoheadrightarrow R, \tag{5.10}$$

such that the composition

$$\bigoplus_{k \in \mathbb{Z}} \mathcal{M}(\mathbb{F}_p, k, \mu_N) \twoheadrightarrow R \hookrightarrow \widehat{R_{\mathbf{Tate}}} \tag{5.11}$$

is the q-expansion map at the cusp **Tate**.

Given the Proposition, if we further prove that $\mathrm{Ker}(r) = (H-1)$ the Theorem is proved.

To construct the map r we first construct modular forms $a(k)$ for $k \in \mathbb{Z}$ with the following properties (compare [**43**]):

- $a(k)$ is a modular form over $\mathbb{F}_p$ of μ_p-level and weight k.
- $a(k)a(k') = a(kk')$
- $a(k)$ does not vanish.

Indeed, given a μ_p level structure

$$\beta_p : \mu_p \hookrightarrow E, \tag{5.12}$$

we have an induced isomorphism

$$d\beta_p : \mathfrak{t}^*_{E/R} \longrightarrow \mathfrak{t}^*_{\mu_p/R} = R \cdot \frac{dt}{t}. \tag{5.13}$$

Thus we get a canonical element $\omega_{can} = (d\beta_p)^{-1}(dt/t)$ in $\mathfrak{t}^*_{E/R}$. This gives the modular form $a(1)$. We note that since ω_{can} is non-vanishing $a(1)$ is non-vanishing.

We simply put $a(k) = a(1)^k$, a non-vanishing modular form of weight k and level μ_p. Moreover, since for the Tate object ω_{can} *is* the canonical differential, the q-expansion of $a(1)$ (and hence of every $a(k)$) is 1.

We now define the map r by

$$\mathcal{M}(\mathbb{F}_p, k, \mu_N) \ni f \mapsto \frac{f}{a(k)} \in R. \tag{5.14}$$

The properties of the forms $a(k)$ guarantee that this provides a well defined ring homomorphism $\underset{k\in\mathbb{Z}}{\oplus}\mathcal{M}(\mathbb{F}_p, k, \mu_N) \longrightarrow R$. Furthermore, since the q-expansion of a modular form is obtained by evaluating it at a Tate object over a base which is isomorphic to the completion of the moduli space at this point, the composition of r with the inclusion $R \hookrightarrow \widehat{R_{\mathbf{Tate}}}$ *is* the q-expansion.

We define an action of $(\mathbb{Z}/p\mathbb{Z})^\times$ on $\underset{k\in\mathbb{Z}}{\oplus}\mathcal{M}(\mathbb{F}_p, k, \mu_N)$ and on R. Given a modular form f of weight k and $\alpha \in (\mathbb{Z}/p\mathbb{Z})^\times$ we define

$$[\alpha]f = \alpha^k f. \tag{5.15}$$

Given a function $g \in R$ we define

$$([\alpha]g)(E, \beta_N \times \beta_p) = g(E, \beta_N \times (\beta_p \circ \alpha)), \tag{5.16}$$

where we denote by $\alpha : \mu_p \longrightarrow \mu_p$ the homomorphism of "raising to the α power".

We claim that the map r is equivariant for this action. Indeed, the same definition given for functions $g \in R$ may well be given for modular forms on $\mathfrak{M}^*(\mu_p)$. Namely, by twisting the μ_p level. Thus

$$[\alpha]r(f) = \frac{[\alpha]f}{[\alpha]a(k)} = \frac{f}{[\alpha]a(k)}. \tag{5.17}$$

(Note: as a modular form on $\mathfrak{M}^*(\mu_p)$ f is invariant under the action of $(\mathbb{Z}/p\mathbb{Z})^\times$. Thus $[\alpha]f = f$. But, as a modular form on $\mathfrak{M}^*$ we have by definition $[\alpha]f = \alpha^k f$). Now,

$$([\alpha]a(1))(E, \beta_p) = a(1)(E, \beta_p \circ \alpha) \tag{5.18}$$

$$= d(\beta_p \circ \alpha)^{-1}(dt/t) \tag{5.19}$$

$$= d\beta_p^{-1} d\alpha^{-1}(dt/t). \tag{5.20}$$

Since $d\alpha(dt/t) = dt^\alpha/t\alpha = \alpha t^{\alpha-1}dt/t^\alpha = \alpha dt/t$, we find that $d\alpha^{-1}(dt/t) = \alpha^{-1}dt/t$ and hence that $[\alpha]a(1) = \alpha^{-1}a(1)$. Therefore,

$$[\alpha]a(k) = \alpha^{-k}a(k), \tag{5.21}$$

and

$$[\alpha]r(f) = \alpha^k r(f) = r([\alpha]f). \tag{5.22}$$

We note that

$$R = \overset{p-2}{\underset{k=0}{\oplus}} R^k, \tag{5.23}$$

where $R^k = \{f \in R : [\alpha]f = \alpha^k f\}$. Then $r(\mathcal{M}(\mathbb{F}_p, k, \mu_N)) \subset R^k$.

We already know that $H - 1 \in \mathrm{Ker}(r)$. Let now $f_1 + \cdots + f_a$ be in the kernel of r. Say, f_i is of weight k_i. We assume that for $i \neq j$, $k_i \neq k_j$. By replacing f_i by

$f_i + f_i(H-1)$ for suitable i's and sufficiently many times, we may assume that if $i \neq j$ then $k_i \neq k_j$ modulo $p-1$. But then every $r(f_i)$ lies in a different summand of the right hand side of Equation (5.23). Hence, every $r(f_i) = 0$.

Note that on any fixed weight the q-expansion , equivalently the map r, is injective. This is evident from Equation (5.14). Thus, for every i, $f_i = 0$.

Finally, the map r is surjective. Let $g \in R^k$. Define $f = g \cdot a(k) \cdot H^n$. Then f is invariant under the Galois group of $\mathfrak{M}^*(\mu_p) \longrightarrow \mathfrak{M}^{\mathrm{ord}}$. Hence, it defines a meromorphic modular form on $\mathfrak{M}$, with poles supported on the supersingular locus. Now, the Hasse invariant vanishes on the supersingular locus. This follows readily from interpreting $H^1_{dR}(E)$ as the Dieudonné module of $E[p]$. See Appendix A. Thus, for $n \gg 0$ the modular form f would be holomorphic. Clearly, $r(f) = g$.

□

Corollary 5.4. *We have*

$$\widetilde{\mathcal{N}}^0 = \bigoplus_{k \equiv 0 \pmod{p-1}} \mathcal{M}(\mathbb{F}, k, \mu_N)/(H-1) \cong R^0; \tag{5.24}$$

R^0 is the ring of regular functions on the affine regular modular curve $\mathfrak{M}(\mathbb{F}, \mu_N)^{\mathrm{ord}}$. Hence, $\widetilde{\mathcal{N}}^0$ is a Dedekind domain.

5.3. Operators revisited. We use the method of Section 5.2 to interprete the operators $T_\ell, < d >, U, V$ and θ via the isomorphism

$$r : \oplus_k \mathcal{M}(\mathbb{F}_p, k, \mu_N)/(H-1) \cong R, \tag{5.25}$$

where R is the ring of regular functions on the affine curve $\mathfrak{M}^*(\mathbb{F}_p, \mu_{Np})$. The a priory observation that the formulae for those operators are given solely in terms of their q-expansion and that for T_ℓ it depends on the weight only modulo $p-1$ indicates that "they are coming from R". To define those operators we shall think of the points of R as pairs (E, β_{Np}).

Let $g \in R$. Define

$$(T_\ell g)(E, \beta_{Np}) = \frac{1}{\ell} \sum_{H \subset E} g(E/H, \pi_H \circ \beta_{Np}); \tag{5.26}$$

The summation ranging over all subgroups of order ℓ of E and $\pi_H : E \longrightarrow E/H$ is the projection. We claim that

$$r(T_\ell f) = T_\ell \, r(f). \tag{5.27}$$

First note that

$$r(f)(E, \beta_{Np}) = f(E, \beta_N, \beta_{p*}(dt/t)). \tag{5.28}$$

where we put $\beta_{p*} = (d\beta_p)^{-1}$. This follows immediately from checking the q-expansions. Hence,

$$r(T_\ell f)(E, \beta_{Np}) = (T_\ell f)(E, \beta_N, \beta_{p*}(dt/t)) \tag{5.29}$$

$$= \frac{1}{\ell} \sum_{H \subset E} f(E/H, \pi_H \circ \beta_N, \omega_H), \tag{5.30}$$

where ω_H is the differential induced from $\beta_{p*}(dt/t)$ via $E \longrightarrow E/H$. That is, $\omega_H = (\pi_H \circ \beta_p)_*(dt/t)$. On the other hand,

$$(T_\ell r(f))(E, \beta_{Np}) = \frac{1}{\ell} \sum_{H \subset E} r(f)(E/H, \pi_H \circ \beta_{Np}) \tag{5.31}$$

$$= \frac{1}{\ell} \sum_{H \subset E} r(f)(E/H, \pi_H \circ \beta_N, (\pi_H \circ \beta_p)_*(dt/t)). \tag{5.32}$$

Formula (5.27) follows.

We defined for a modular form f

$$(< d > f)(E, \beta_N, \omega) = f(E, \beta_N \circ d, \omega). \tag{5.33}$$

We define for $g \in R$

$$(< d > g)(E, \beta_{Np}) = (E, \beta_N \circ d, \beta_{p*}(dt/t)). \tag{5.34}$$

Clearly,

$$r(< d > f) =< d > r(f). \tag{5.35}$$

We defined for a modular form f

$$(Vf)(E, \beta_N, \omega) = f(E^{(p)}, \beta_N^{(p)}, \omega^{(p)}). \tag{5.36}$$

The scheme $\mathfrak{M}(\mathbb{F}, \mu_{Np})$ is a scheme defined over $\mathbb{F}_p$. We thus have a Frobenius morphism

$$\mathrm{Fr} : \mathfrak{M}(\mathbb{F}, \mu_{Np}) \longrightarrow \mathfrak{M}(\mathbb{F}, \mu_{Np}). \tag{5.37}$$

Given a function $g \in R$ we let

$$Vg = g \circ \mathrm{Fr}. \tag{5.38}$$

From a moduli perspective, since the point corresponding to $(E^{(p)}, \beta_{Np}^{(p)})$ is Fr of the point corresponding to (E, β_{Np}):

$$(Vg)(E, \beta_{Np}) = g(E^{(p)}, \beta_{Np}^{(p)}). \tag{5.39}$$

In particular,

$$(Vr(f))(E, \beta_{Np}) = f(E^{(p)}, \beta_N^{(p)}, (\beta_p^{(p)})_*(dt/t)) \tag{5.40}$$

$$= f(E^{(p)}, \beta_N^{(p)}, (\beta_{p*}(dt/t))^{(p)}) \tag{5.41}$$

$$= (Vf)(E, \beta_N, \beta_{p*}(dt/t)) \tag{5.42}$$

$$= r(Vf)(E, \beta_{Np}). \tag{5.43}$$

That is,

$$r(Vf) = Vr(f). \tag{5.44}$$

Given an elliptic curve E let $E[V]$ denote the kernel of Verschiebung. Given a pair (E, β_N) (resp. (E, β_{Np})) we get a well-defined pair $(E/E[V], \pi_{E[v]} \circ \beta_N)$ (resp. $(E/E[V], \pi_{E[v]} \circ \beta_{Np})$). We defined for a modular form f

$$(Uf)(E, \beta_N, \omega) = f(E/E[V], \pi_{E[v]} \circ \beta_N, \omega_{E[V]}), \tag{5.45}$$

where $\omega_{E[V]}$ is the differential on $E/E[V]$ which is the image of ω under the separable map $\pi_{E[v]} : E \longrightarrow E/E[V]$. We define for a function $g \in R$

$$(Ug)(E, \beta_{Np}) = g(E/E[V], \pi_{E[v]} \circ \beta_{Np}). \tag{5.46}$$

One easily verifies

$$r(Uf) = Ur(f). \tag{5.47}$$

Finally, given a function $g \in R$ we define

$$\theta g = \frac{dg}{KS(a(2))}. \tag{5.48}$$

We explain our notation. There is a well known isomorphism, called the Kodaira-Spencer isomorphism,

$$KS : (\mathfrak{t}^*_{\mathcal{E}^U/\mathfrak{M}(\mathbb{F},\mu_N)})^{\otimes 2} \longrightarrow \Omega^1_{\mathfrak{M}(\mathbb{F},\mu_N)/\mathbb{F}}(\mathbf{cusps}), \tag{5.49}$$

for every field k of characteristic prime to N. In particular, one can associate to a modular form f of weight 2 over $\mathfrak{M}(\mathbb{F}, \mu_N)$ a differential $KS(f)$ on that curve. This differential is holomorphic iff the modular form is a cusp form.

Over the complex numbers this is very familiar: If f is a modular form of weight 2 (i.e. a section of $(\mathfrak{t}^*_{\mathcal{E}^U/\mathfrak{M}(\mathbb{C},\mu_N)})^{\otimes 2}$) then $f(\tau)d\tau$ is a meromorphic differential with at most simple poles supported at the cusps. Indeed

$$f(\gamma\tau)d\gamma\tau = (c\tau + d)^2 f(\tau) \cdot \frac{(c\tau + d)a - (a\tau + b)c}{(c\tau + d)^2} d\tau \tag{5.50}$$

$$= f(\tau)d\tau. \tag{5.51}$$

At $i\infty$, if $f(q) = \sum_n a_n q^n$ then $q = \exp(2\pi i \cdot \tau)$ implies $dq/q = 2\pi i \cdot d\tau$ and therefore $f(\tau)d\tau = 2\pi i \cdot f(q)(dq/q)$, which is holomorphic if and only if f is a cusp form.

The covering

$$\mathfrak{M}^*(\mathbb{F}_p, \mu_{Np}) \longrightarrow \mathfrak{M}^*(\mathbb{F}_p, \mu_N)^{\mathrm{ord}}, \tag{5.52}$$

of non-singular affine curves, extends uniquely to a covering of non-singular proper curves

$$\mathfrak{M}^\dagger(\mathbb{F}_p, \mu_{Np}) \longrightarrow \mathfrak{M}^*(\mathbb{F}_p, \mu_N). \tag{5.53}$$

Moreover, $\mathfrak{M}^\dagger(\mathbb{F}_p, \mu_{Np})$ may be defined intrinsically as $\mathfrak{M}^*(\mathbb{F}_p, \mu_N)[a(1)]$ – the scheme obtained from $\mathfrak{M}^*(\mathbb{F}_p, \mu_N)$ by adjoining the $(p-1)$-st root of the section H of the line bundle $(\mathfrak{t}^*_{\mathcal{E}^U/derM^*(\mathbb{F}_p,\mu_{Np})})^{\otimes p-1}$.

The morphism $\mathfrak{M}^\dagger(\mathbb{F}_p, \mu_{Np}) \longrightarrow \mathfrak{M}^*(\mathbb{F}_p, \mu_N)$ is thus a finite separable morphism of degree $p-1$, commuting with the action of $(\mathbb{Z}/p\mathbb{Z})^\times$ (in fact, $\mathfrak{M}^*(\mathbb{F}_p, \mu_N)$ is the *quotient* by this action), and its ramification divisor is $(p-2)W_1$, where W_1 is the supersingular locus. In particular, it is totally ramified over the supersingular locus.

Now, when dealing with level μ_{Np}, $(N, p) = 1$, in characteristic p, the Kodaira-Spencer isomorphism needs to be modified. On the one hand, if we let f denote the morphism $\mathfrak{M}^\dagger(\mathbb{F}_p, \mu_{Np}) \longrightarrow \mathfrak{M}^*(\mathbb{F}_p, \mu_N)$ then

$$f^*(\mathfrak{t}^*_{\mathcal{E}^U/\mathfrak{M}(\mathbb{F},\mu_N)}) = \mathfrak{t}^*_{\mathcal{E}^U/\mathfrak{M}(\mathbb{F},\mu_{Np})}. \tag{5.54}$$

On the other, since the separable morphism $f : \mathfrak{M}^*(\mathbb{F}_p, \mu_{Np}) \longrightarrow \mathfrak{M}(\mathbb{F}_p, \mu_N)$ is completely (but tamely) ramified over the supersingular locus W_1, we get

$$f^*\Omega^1_{\mathfrak{M}^*(\mathbb{F}_p,\mu_N)} = \Omega^1_{\mathfrak{M}^\dagger(\mathbb{F}_p,\mu_{Np})}((2-p)W_1). \tag{5.55}$$

(See [**47**, IV.2]). Putting it all together, we get an isomorphism

$$KS : (\pi_* \mathfrak{t}^*_{E^U/\mathfrak{M}(\mathbb{F},\mu_{Np})})^{\otimes 2} \longrightarrow \Omega^1_{\mathfrak{M}^*(\mathbb{F}_p,\mu_{Np})}(\mathbf{cusps} + (2-p)W_1). \tag{5.56}$$

It follows that $a(2)$ defines a differential whose order of vanishing along W_1 is $p = 2 + (p-2)$.

We claim that $r(\theta f) = \theta r(f)$. This is readily checked via the q-expansions. If the q-expansion of a function g is $\sum a_n q^n$ then θg has expansion $(\sum n a_n q^{n-1} dq)/(dq/q)$ (the differential on the Tate curve corresponding to $a(2)$ is dq/q). Thus,

$$r(\theta f) = \theta r(f). \tag{5.57}$$

Some of the delicate behaviour of the operator θ can be explained by this interpretation. First

LEMMA 5.5. *Let $g \in R^{k_0}$. Then the expansion of g around every supersingular point P has the form*

$$a_{k_P} x^{k_P} + a_{k_P+p-1} x^{k_P+p-1} + a_{k_P+2(p-1)} x^{k_P+2(p-1)} + \dots, \tag{5.58}$$

where $k_P \equiv k_0 \pmod{p-1}$. The filtration of the q-expansion $g(q)$ is

$$w(g(q)) = -\min\{k_P : P \in W_1\}. \tag{5.59}$$

PROOF. The shape of the Taylor expansion follows immediately from the fact that every supersingular point is a fixed point of the group action $(\mathbb{Z}/p\mathbb{Z})^\times$ and $g \in R^{k_0}$. The second fact is also evident, because $a(b)$, as a modular form, vanishes to order b at such point. □

COROLLARY 5.6. *Let f be a modular form of weight k and filtration $w(f)$. Then:*

1. *$w(f) < k$ iff f vanishes along W_1, or equivalently, $H|f$;*
2. *$w(Vf) = p \cdot w(f)$;*
3. *$w(\theta f) \leq w(f) + p + 1$ with equality holding iff $p \not| w(f)$.*

PROOF. The first assertion is immediate. Note that the Taylor expansion of $r(f)$ at a supersingular point P starts with x^{-k+t} where t is not zero iff f vanishes at P.

To see the effect of V, we note that the Taylor expansion of $r(f)$ at some supersingular point starts with $a_{-w(f)} x^{-w(f)}$ ($a_{-w(f)} \neq 0$) and at all other points with terms of not smaller degree. The Taylor expansion of $Vr(f)$ starts at some point with $cx^{-pw(f)}$, because $Vr(f) = r(f) \circ \mathrm{Fr}$ and at all other points with terms of not smaller degree.

To prove the last assertion, note that if at some supersingular point $r(f)$ has expansion $a_{k_P} x^{k_P} + a_{k_P+p-1} x^{k_P+p-1} + a_{k_P+2(p-1)} x^{k_P+2(p-1)} + \dots$ then $\theta r(f)$ has expansion starting with ckx^{k-1-p} (using that dg starts with kx^{k-1} and $a(2)$ with x^{-p}). Thus, $w(\theta f) \leq w(f) + p + 1$ with equality iff $p \not| w(f)$. □

6. p-adic Elliptic Modular Forms

6.1. Test objects and overconvergent forms. Let B be a p-adic ring. By that we mean that $B \cong \varprojlim B/p^nB$, and so, in particular, B is a $\mathbb{Z}_p$-algebra. Note that examples are provided by the ring of integers R of a finite extension of $\mathbb{Q}_p$, and by any quotient $R/(\pi^n)$ of R, where π is a uniformizer. One may take $n = 1$, or more generally any field of characteristic p.

We fix:

- B – a p-adic ring. ("The base ring").
- r – an element of B. ("The growth gauge").

Definition 6.1. Let A be a p-adic ring which is a B-algebra. A *test object* of:

- level μ_N,
- growth condition r,
- over B,

is a quadruple:

$$(E, \omega, \beta_N, Y)_{/A}, \tag{6.1}$$

such that:

- E/A is an elliptic curve, i.e. a proper, smooth group scheme over $\mathrm{Spec}(A)$, such that the geometric fibers are connected curves of genus one;
- $\omega \in \mathfrak{t}^*_{E/A}$ is a relative non-vanishing differential;
- $\beta_N : \mu_{N/A} \hookrightarrow E$ is an embedding of group schemes over $\mathrm{Spec}(A)$;
- $Y \in A$ satisfies $Y \cdot E_{p-1}(E, \omega) = r$.

Example 6.2. 1. Given $(E, \omega, \beta_N)_{/A}$, such Y need not exist. For example: if $r = 1$ (or any unit) the condition is that $E_{p-1}(E, \omega) \in A^\times$. This excludes every elliptic curve E such that $E \pmod{\mathfrak{m}}$ is supersingular, where $\mathfrak{m}$ is a maximal ideal. Indeed: the ring A is p-adic, so $pA \neq A$, and we take $\mathfrak{m}$ to be any maximal ideal. Note that if $a \in pA$ then $(1-a)^{-1} = 1 + a + a^2 + \ldots$ converges in A. Therefore pA is contained in the Jacobson radical of A. That is, every maximal ideal contains pA. The reduction $(E \mod \mathfrak{m}, \omega \mod \mathfrak{m})$ is an elliptic curve together with a non-vanishing differential. Since the reduction modulo p of E_{p-1} is the Hasse invariant, it vanishes at all supersingular elliptic curves over fields of characteristic p with any choice of differential. Thus

$$E_{p-1}(E, \omega) \pmod{\mathfrak{m}} \equiv E_{p-1}(E \mod \mathfrak{m}, \omega \mod \mathfrak{m}) = 0 \mod \mathfrak{m} \quad \text{iff } E \pmod{\mathfrak{m}} \text{ is supersingular.} \tag{6.2}$$

2. Examining the idea further, note that if $E_{p-1}(E, \omega) = p$, then indeed it may happen that $p^{1/(p-1)} \in A$ and then $E_{p-1}(E, p^{1/(p-1)}\omega) = 1$, but $p^{\frac{1}{p-1}}\omega$ is not a non-vanishing differential <u>over A</u>. That is $(E, p^{1/(p-1)}\omega)$ is <u>not</u> a test object over A.
3. On the other extreme, if $r = p$, and say $E_{p-1}(E, \omega) = p$, then $(E, \omega, \beta_N, 1)$ is a test object.
4. Given $(E, \omega, \beta_N, Y)_{/A}$, any other Y is of the form $Y + t$, with $t \cdot E_{p-1}(E, \omega) = 0$. Note that *if $E_{p-1}(E, \omega)$ is invertible, Y is uniquely determined.*

5. If $(E,\omega,\beta_N,Y)_{/A}$ is a test object, so is $(E,\lambda\omega,\beta_N,\lambda^{p-1}Y)_{/A}$ for $\lambda\in A^\times$. Indeed,

$$(\lambda^{p-1}Y)\cdot(E_{p-1}(E,\lambda\omega))=\lambda^{p-1}Y\lambda^{-(p-1)}E_{p-1}(E,\omega)=r. \tag{6.3}$$

Equivalently, a test object could be thought of as that data

$$(E,\beta_N,Y)_{/A},\ \ Y\in(\mathfrak{t}^*_{E/A})^{-(p-1)} \tag{6.4}$$

such that $Y\cdot E_{p-1}=r$, where E_{p-1} is interpreted as a rule associating to E/A a section of $(\mathfrak{t}^*_{E/A})^{p-1}$.

Definition 6.3. A *p-adic modular form* (à la Katz) of:

- weight $k\in\mathbb{Z}$,
- level μ_N,
- growth r,
- defined over B,

is a rule associating to a test object (of level μ_N and growth r) $(E,\omega,\beta_N,Y)_{/A}$ an element $f(E,\omega,\beta_N,Y)\in A$ such that:

- $f(E,\omega,\beta_N,Y)$ depends only on the isomorphism class of the test object $(E,\omega,\beta_N,Y)_{/A}$;
- the rule f commutes with base change;
- for every $\lambda\in A^\times$ and every test object $(E,\omega,\beta_N,Y)_{/A}$

$$f(E,\lambda\omega,\beta_N,\lambda^{p-1}Y)_{/A}=\lambda^{-k}f(E,\omega,\beta_N,Y)_{/A}. \tag{6.5}$$

Definition 6.4. The space of modular forms over B, of μ_N-level, weight k and growth condition r is denoted by $\mathbb{F}(B,k,\mu_N;r)$. If $r\notin B^\times$, they are called *overconvergent* modular forms.

Given $s\in B$, we have a B-module homomorphism

$$\mathbb{F}(B,k,\mu_N;sr)\longrightarrow\mathbb{F}(B,k,\mu_N;r), \tag{6.6}$$

$$f\mapsto f',\ \ f'(E,\omega,\beta_N,Y)_{/A}=f(E,\omega,\beta_N,sY)_{/A}. \tag{6.7}$$

If $s\in B^\times$, then this is an isomorphism. Thus, whenever $r\in B^\times$, we might as well take $r=1$. We shall see in Corollary 6.15 that for $r=1$ these are Serre's p-adic modular forms.

In the case $r=1$, as we pointed out, we discard all elliptic curves with supersingular reduction; but if we consider growth condition $r\notin B^\times$, it could very well happen that the modular form in question is defined also on some portion of the supersingular disks of the moduli space of elliptic curves with level μ_N-structure, besides being defined on the complement. This is why such forms are called "overconvergent". To be precise, the very definition of the supersingular disks uses E_{p-1} ([**14**, Sections 1-2]). The disk of radius n may be defined as all the $\overline{\mathbb{Q}_p}$ points x of $\mathfrak{M}(\overline{\mathbb{Q}_p},\mu_N)$ such that E_x has good reduction modulo p and for one (equivalently, any) differential ω_x on E_x such that (E_x,ω_x) has good reduction (i.e., ω_x extends to a non-vanishing differential over $\mathcal{O}_{\overline{\mathbb{Q}_p}}$) we have $|E_{p-1}(E_x,\omega_x)|\leq p^{-n}$. In fact the growth r actually controls the growth of the coefficients of the Laurent expansions of these forms around the supersingular disks.

Every classical modular form f of weight k, level N, over B defines a p-adic modular form (still denoted f) in $\mathbb{F}(B,k,\mu_N;r)$:

$$f(E,\omega,\beta_N,Y)_{/A} := f(E,\omega,\beta_N)_{/A}. \tag{6.8}$$

As an example of a truly p-adic modular form, consider the modular form in $\mathbb{F}(B,1-p,\mu_N;r)$ given by:

$$f(E,\omega,\beta_N,Y)_{/A} = Y. \tag{6.9}$$

This is indeed a p-adic modular form. The requirement on base change and isomorphism class being obvious. By definition,

$$f(E,\lambda\omega,\beta_N,\lambda^{p-1}Y) = \lambda^{p-1}Y = \lambda^{-(1-p)}f(E,\omega,\beta_N,Y). \tag{6.10}$$

Therefore, the weight is $1-p$.

If $r=1$, then $Y\cdot E_{p-1}(E,\omega)=1$, i.e.,

$$Y = \frac{1}{E_{p-1}}. \tag{6.11}$$

We remark that this is also a modular form à la Serre (Exercise 4.4).

6.2. q-expansion for p-adic modular forms. If $E_{p-1}(E,\omega)$ is invertible, then in any test object (E,ω,β_N,Y) we must take $Y = r\cdot E_{p-1}(E,\omega)^{-1}$. This applies in particular to the Tate curve $\mathbf{Tate}(q)=\mathbb{G}_m/q(\mathbb{Z})$ over B. The Tate curve carries a canonical μ_N-level:

$$\beta_{can} : \mu_N \hookrightarrow \mathbb{G}_m \longrightarrow \mathbf{Tate}(q), \tag{6.12}$$

and a canonical differential ω_{can} induced from the differential dt/t on $\mathbb{G}_m$. The cusp $(\mathbf{Tate}(q),\beta_{can})$ is called the standard cusp.

Definition 6.5. Let f be a p-adic modular form,$f\in\mathbb{F}(B,k,\mu_N;r)$. The q-expansion of f is

$$f\left(\mathbf{Tate}(q),\omega_{can},\beta_{can},\frac{r}{E_{p-1}(\mathbf{Tate}(q),\omega_{can})}\right)\in B((q)). \tag{6.13}$$

Similarly, for any cusp $(\mathbf{Tate}(q),\beta_N)$ we define the q-expansion of f by

$$f\left(\mathbf{Tate}(q),\omega_{can},\beta_N,\frac{r}{E_{p-1}(\mathbf{Tate}(q),\omega_{can})}\right)\in B((q)). \tag{6.14}$$

We call f *holomorphic* (respectively, *cusp*) form if all its q-expansion lies in $B[[q]]$ (respectively, in $q\cdot B[[q]]$ at every cusp). We denote the holomorphic (resp. cusp) forms by

$$\mathbb{M}(B,k,\mu_N;r) \quad (\text{resp. } \mathbb{S}(B,k,\mu_N;r)). \tag{6.15}$$

Proposition 6.6. *Take any* $\mathbb{X}\in\{\mathbb{F},\mathbb{M},\mathbb{S}\}$. *Then*

$$\mathbb{X}(B,k,\mu_N;r) = \varprojlim \mathbb{X}(B/p^nB,k,\mu_N;r). \tag{6.16}$$

Proof. All our objects are p-adic: $B=\varprojlim B/p^nB$, and to give an elliptic curve over A is equivalent to giving an inductive system

$$E_{1_{/(A/pA)}} \hookrightarrow E_{2_{/(A/p^2A)}} \hookrightarrow E_{3_{/(A/p^3A)}} \hookrightarrow \dots \ . \tag{6.17}$$

□

We offer some remarks:

1. The q-expansion of a p-adic modular form f' induced from a classical modular form f as in (6.8) is the same q-expansion as of f.
2. The q-expansion is injective. See Corollary 6.13.

6.3. The case when p is nilpotent. In this section we let $p \geq 5$ be a prime and $N \geq 4$ is an integer prime to p. As before B be a p-adic ring. We now assume further that p is nilpotent in B.

We use our usual notation

$$\mathcal{M}(B, k, \mu_N) \tag{6.18}$$

for classical holomorphic modular forms over B, of μ_N-level structure and weight k. We denote by

$$\mathcal{F}(B, k, \mu_N) \tag{6.19}$$

the classical modular forms over B, of μ_N-level structure and weight k with possible poles at infinity. As above, $\mathbb{F}(B, k, \mu_N; r)$ is the space of p-adic modular forms with growth r. Given $j \geq 0$, define a map:

$$\mathcal{F}(B, k + j(p-1), \mu_N) \longrightarrow \mathbb{F}(B, k, \mu_N; r) \tag{6.20}$$

$$f \mapsto \widetilde{f}, \quad \widetilde{f}(E, \omega, \beta_N, Y)_{/A} = Y^j f(E, \omega, \beta_N). \tag{6.21}$$

We claim that $\widetilde{f}$ is a p-adic modular form of weight k. Indeed

$$\begin{aligned} \widetilde{f}(E, \lambda\omega, \beta_N, \lambda^{p-1}Y) &= (\lambda^{p-1}Y)^j f(E, \lambda\omega, \beta_N) \\ &= \lambda^{j(p-1)} Y^j \lambda^{-(k+j(p-1))} f(E, \omega, \beta_N) \\ &= \lambda^{-k} \widetilde{f}(E, \omega, \beta_N, Y). \end{aligned} \tag{6.22}$$

(Alternately, Y^j is a modular form of weight $-j(p-1)$. See (6.10)). Under this map, $E_{p-1}f$ is sent to $\widetilde{E_{p-1}f}$, and

$$\begin{aligned} \widetilde{E_{p-1}f}(E, \omega, \beta_N, Y) &= Y^{j+1} f(E, \omega, \beta_N) \cdot E_{p-1}(E, \omega, \beta_N) \\ &= r \cdot Y^j f(E, \omega, \beta_N) \\ &= r \cdot \widetilde{f}. \end{aligned} \tag{6.23}$$

Therefore, we have obtained a well-defined homomorphism of B-modules as follows:

$$\left(\bigoplus_{j \geq 0} \mathcal{F}(B, k + j(p-1), \mu_N) \right) / (E_{p-1} - r) \longrightarrow \mathbb{F}(B, k, \mu_N; r). \tag{6.24}$$

Here $(E_{p-1} - r)$ stands for the submodule generated by $\{(E_{p-1} - r)f : f \in \mathcal{F}(B, k + j(p-1), \mu_N)\}$.

Proposition 6.7. *The map*

$$\left(\bigoplus_{j \geq 0} \mathcal{F}(B, k + j(p-1), \mu_N) \right) / (E_{p-1} - r) \xrightarrow{\cong} \mathbb{F}(B, k, \mu_N; r) \tag{6.25}$$

is an isomorphism.

PROOF. Consider the functors:

$\underline{\underline{F_1}}$:

$$A \in B\text{-algebra} \ \mapsto F_1(A), \tag{6.26}$$

where $F_1(A)$ consists of isomorphism classes of the following data:

- $(E, \beta_N, Y)_{/A}$,
- $Y \in \underline{\omega}_E^{-(p-1)}$, such that $Y \cdot E_{p-1} = r$,

where $\underline{\omega}_E = \mathfrak{t}^*_{E/A}$.

This functor is equivalent to the functor associating to $A \in B$-algebra the following data:

- $g : \mathrm{Spec}(A) \longrightarrow \mathfrak{M}(B, \mu_N)$,
- $Y \in g^*(\mathcal{L})$, such that $Y \cdot g^* E_{p-1} = r$,

where $\mathcal{L} = \underline{\omega}^{-(p-1)}, \underline{\omega} = \mathfrak{t}^*_{\mathcal{E}^U/\mathfrak{M}(B,\mu_N)}$.

The functor $F1$ is a subfunctor of $F2$:

$\underline{\underline{F_2}}$:

$$A \in B\text{-algebra} \ \mapsto F_2(A), \tag{6.27}$$

where $F_2(A)$ consists of isomorphism classes of the following data:

- $(E, \beta_N, Y)_{/A}$,
- $Y \in \underline{\omega}_E^{-(p-1)}$,

which is by the same token equivalent to the functor associating to A the following:

- $g : \mathrm{Spec}(A) \longrightarrow \mathfrak{M}(B, \mu_N)$,
- $Y \in g^*(\mathcal{L})$.

The last functor is representable by the $\mathfrak{M}(B, \mu_N)$-scheme $\mathbf{Spec}(S(\mathcal{L}^\vee))$.[7] The condition that a section Y satisfies $Y \cdot g^*(E_{p-1}) = r$ is exactly that locally, if the homomorphism $\mathbf{Spec}(\mathbf{S}(\mathcal{L}^\vee)) \longrightarrow \mathcal{M}(\mathbf{B}, \mu_{\mathbf{N}})$ describing Y is given by $A_i[x_1] \xrightarrow{Y} A_i$, we have $Y(E_{p-1}) = Y \cdot E_{p-1} = r$.

Hence F_1 is representable by the $\mathfrak{M}(B, \mu_N)$ scheme:

$$\alpha = \mathbf{Spec}(S(\underline{\omega}^{p-1})/(E_{p-1} - r)) \hookrightarrow \mathbf{Spec}(S(\underline{\omega}^{p-1})) \longrightarrow \mathfrak{M}(B, \mu_N). \tag{6.28}$$

We write $f : \alpha \longrightarrow \mathfrak{M}(B, \mu_N)$.

[7] The notation is as follows: for a line bundle $\mathcal{L} \longrightarrow X$ we let $\mathcal{L}^\vee = \mathcal{L}^{-1}$ denote the dual line bundle on X. We let $S(\mathcal{L})$ denote the sheaf of symmetric algebra on X constructed from $\mathcal{L}$. We let $\mathbf{Spec}(S(\mathcal{L}^\vee)) \longrightarrow X$ denote the associated affine morphism. See [**47**, Exercises II 5.16-5.18].

We may identify the space $\mathbb{F}(B,k,\mu_N;r)$ of p-adic modular forms over B of weight k with growth r as follows:

(6.29)

$$\mathbb{F}(B,k,\mu_N;r) = H^0(\alpha, f^*\underline{\omega}^k)$$

$$= H^0(\mathfrak{M}(B,\mu_N), \underline{\omega}^k \otimes f_*\mathcal{O}_\alpha) \text{ (Leray's spectral seq., [47, Ex. III, 4.1])} \tag{6.30}$$

$$= H^0\left(\mathfrak{M}(B,\mu_N), \underline{\omega}^k \otimes \oplus_{j=0}^\infty \underline{\omega}^{j(p-1)}/(E_{p-1}-r)\right) \tag{6.31}$$

$$= H^0\left(\mathfrak{M}(B,\mu_N), \oplus_{j=0}^\infty \underline{\omega}^{k+j(p-1)}/(E_{p-1}-r)\right) \tag{6.32}$$

$$=^* H^0\left(\mathfrak{M}(B,\mu_N), \oplus_{j=0}^\infty \underline{\omega}^{k+j(p-1)}\right)/(E_{p-1}-r) \tag{6.33}$$

$$= \oplus_{j=0}^\infty \mathcal{F}(B,k+j(p-1),N)/(E_{p-1}-r). \tag{6.34}$$

Let us explain the equality marked by $*$. Note that $b = \oplus_{j=0}^\infty \underline{\omega}^{k+j(p-1)}$ is quasi-coherent sheaf on the *affine* scheme $\mathfrak{M}(B,\mu_N)$. We have an exact sequence:

$$0 \longrightarrow b \xrightarrow{\times(E_{p-1}-r)} b \longrightarrow b/(E_{p-1}-1) \longrightarrow 0\,. \tag{6.35}$$

Since all higher cohomology of quasi-coherent sheaves on affine schemes vanish we get an exact sequence:

$$0 \longrightarrow H^0(\mathfrak{M}(B,\mu_N),b) \overset{\times(E_{p-1}-r)}{\longrightarrow} H^0(\mathfrak{M}(B,\mu_N),b) \longrightarrow H^0(\mathfrak{M}(B,\mu_N), b/(E_{p-1}-r)) \longrightarrow 0\,. \tag{6.36}$$

□

A similar argument, using the compactified curve $\mathfrak{M}^*(B,\mu_N)$ gives

$$\mathbb{M}(B,k,\mu_N;r) = H^0\left(\mathfrak{M}^*(B,\mu_N), (\oplus_{j=0}^\infty \underline{\omega}^{k+j(p-1)})/(E_{p-1}-r)\right). \tag{6.37}$$

Though $(\oplus_{j=0}^\infty \mathcal{M}(B,k+j(p-1),\mu_N))/(E_{p-1}-r)$ is contained in this space, generally they are not equal. However, when r is a p-adic unit they are.

Proposition 6.8. *Let r be a p-adic unit. Then*

$$\mathbb{M}(B,k,\mu_N;r) \cong \left(\oplus_{j=0}^\infty \mathcal{M}(B,k+j(p-1),\mu_N)\right)/(E_{p-1}-r), \tag{6.38}$$

and

$$\mathbb{S}(B,k,\mu_N;r) \cong \left(\oplus_{j=0}^\infty \mathcal{S}(B,k+j(p-1),\mu_N)\right)/(E_{p-1}-r). \tag{6.39}$$

Proof. The idea is to use the affine curve

$$\mathfrak{M}(B,\mu_N)^{\mathrm{ord}} := \mathfrak{M}^*(B,\mu_N) - \{\text{supersingular locus}\} \tag{6.40}$$

instead of the affine curve $\mathfrak{M}(B,\mu_N)$ used before. We may assume without loss of generality that $r=1$. Consider:

$\underline{\underline{F_1}}$: The functor associating to a B-algebra A the data:

- $g : \mathrm{Spec}(A) \longrightarrow \mathfrak{M}^*(B,\mu_N)$,
- $Y \in g^*(\mathcal{L})$ such that $Y \cdot g^*E_{p-1} = 1$,

where $\mathcal{L} = \underline{\omega}^{-(p-1)}$.

$\underline{\underline{F_2}}$: The functor associating to a B-algebra A the data:

- $g : \mathrm{Spec}(A) \longrightarrow \mathfrak{M}^*(B, \mu_N)$,
- $Y \in g^*(\mathcal{L})$.

The functor F_1 is a subfunctor of the functor F_2, which is represented by the scheme

$$\mathbf{Spec}(S(\mathcal{L}^\vee)) \longrightarrow \mathfrak{M}^*(B, \mu_N). \tag{6.41}$$

A similar reasoning gives that F_1 is represented by the scheme

$$\beta = \mathbf{Spec}(S(\mathcal{L}^\vee)/(E_{p-1} - 1)) \longrightarrow \mathfrak{M}^*(B, \mu_N). \tag{6.42}$$

Still better, the structural morphism of β factors through $\mathfrak{M}(B, \mu_N)^{\mathrm{ord}}$. Thus F_1 is represented by

$$f : \beta \longrightarrow \mathfrak{M}(B, \mu_N)^{\mathrm{ord}}. \tag{6.43}$$

Let $\underline{\omega}$ denote the sheaf of modular forms on $\mathfrak{M}^{\mathrm{ord}}(B, \mu_N)$. Let $\mathbb{X}$ be one of the symbols $\mathbb{M}, \mathbb{S}$. We define a sheaf b by

$$\mathbb{X} = \begin{cases} \underline{\omega}^k & \mathbb{X} = \mathbb{M} \\ \underline{\omega}^k(-\mathbf{cusps}) & \mathbb{X} = \mathbb{S} \end{cases}. \tag{6.44}$$

Then

$$\mathbb{X}(B, k, \mu_N, 1) = H^0(\beta, b) \tag{6.45}$$

$$= H^0(\mathfrak{M}(B, \mu_N)^{\mathrm{ord}}, b \otimes f_* \mathcal{O}_\beta) \tag{6.46}$$

$$= H^0\left(\mathfrak{M}(B, \mu_N)^{\mathrm{ord}}, b \otimes (\oplus_{j=0}^\infty \underline{\omega}^{j(p-1)})/(E_{p-1} - 1)\right) \tag{6.47}$$

$$= H^0\left(\mathfrak{M}(B, \mu_N)^{\mathrm{ord}}, b \otimes \oplus_{j=0}^\infty \underline{\omega}^{j(p-1)}\right)/(E_{p-1} - 1) \tag{6.48}$$

$$= \oplus_{j=0}^\infty X(B, k + j(p-1), \mu_N)/(E_{p-1} - 1). \tag{6.49}$$

Here $X = \mathcal{M}$ (resp. $\mathcal{S}$) if $\mathbb{X} = \mathbb{M}$ (resp. $\mathbb{S}$), and the equalities follows from the same considerations as in the proof of Proposition 6.7. □

6.4. The case of r a unit. An immediate consequence of the Propositions 6.7 and 6.8 is the following

THEOREM 6.9. *Let $N \geq 4$ be an integer. Let $p \geq 5$ be prime and let B be a p-adic ring. Let $r \in B^\times$. Then*

$$\mathbb{M}(B, k, \mu_N; r) = \varprojlim_n \left(\oplus_{j=0}^\infty \mathcal{M}(B/p^n B, k + j(p-1), \mu_N)\right)/(E_{p-1} - r), \tag{6.50}$$

and

$$\mathbb{S}(B, k, \mu_N; r) = \varprojlim_n \left(\oplus_{j=0}^\infty \mathcal{S}(B/p^n B, k + j(p-1), \mu_N)\right)/(E_{p-1} - r). \tag{6.51}$$

Note that we could have equally written

$$\mathbb{M}(B, k, \mu_N; r) = \varprojlim_n \left(\oplus_{j=0}^\infty \mathcal{M}(\mathbb{Z}, k + j(p-1), \mu_N) \otimes B/p^n B\right)/(E_{p-1} - r), \tag{6.52}$$

and

$$\mathbb{S}(B, k, \mu_N; r) = \varprojlim_n \left(\oplus_{j=0}^\infty \mathcal{S}(\mathbb{Z}, k + j(p-1), \mu_N) \otimes B/p^n B\right)/(E_{p-1} - r). \tag{6.53}$$

Note that the sums appearing on the right hand sides are not finite anymore. To gain some firm grip on the space of p-adic modular form we introduce a kind of "basis" for that space. This is the topic of the next section.

6.5. Katz's expansion. Consider the map:

$$\mathcal{M}(\mathbb{Z}_p, k+j(p-1), \mu_N) \xrightarrow{\times E_{p-1}} \mathcal{M}(\mathbb{Z}_p, k+(j+1)(p-1), \mu_N). \tag{6.54}$$

We note that upon reduction modulo p this map is *injective*, as is obvious from looking at q-expansions. This implies:

- The map in (6.54) is injective.
- The map in (6.54) splits.

We choose complements:

$$\begin{aligned}\mathcal{M}(\mathbb{Z}_p, k+(j+1)(p-1), \mu_N) = \\ E_{p-1}\cdot\mathcal{M}(\mathbb{Z}_p, k+j(p-1), \mu_N) \ \oplus \ A(\mathbb{Z}_p, k+(j+1)(p-1), \mu_N).\end{aligned} \tag{6.55}$$

($A(\mathbb{Z}_p, k, \mu_N) = \mathcal{M}(\mathbb{Z}_p, k, \mu_N)$). We may tensor with B and we get the same equality with B-coefficients. Then

$$\bigoplus_{j=0}^{j} A(B, k+a(p-1), \mu_N) \cong \mathcal{M}(B, k+j(p-1), \mu_N); \tag{6.56}$$

the map given by

$$(f_0, \dots, f_j) \mapsto \sum_{a=0}^{j} f_a \cdot E_{p-1}^{j-a}. \tag{6.57}$$

Consider the p-adically complete B-module:

$$\begin{aligned}A^{\mathrm{rigid}}(B, k, \mu_N) = \\ \left\{\sum_{a=0}^{\infty} b_a : b_a \in A(B, k+a(p-1), \mu_N),\ b_a \longrightarrow 0 \text{ } p\text{-adically uniformly}\right\}.\end{aligned} \tag{6.58}$$

(i.e. $\forall n, \exists c(n)$ such that $a > c(n)$ implies $p^n | b_a$). If B is a d.v.r. with quotient field K then taking $A^{\mathrm{rigid}}(B, k, \mu_N)$ to be the unit ball in $A^{\mathrm{rigid}}(K, k, \mu_N)$ we get that $A^{\mathrm{rigid}}(K, k, \mu_N)$ is a p-adic Banach space.

Proposition 6.10. (Katz's expansion) *For every growth condition r there exists an isomorphism:*

$$A^{\mathrm{rigid}}(B, k, \mu_N) \xrightarrow[\psi]{\sim} \mathbb{M}(B, k, \mu_N; r) \tag{6.59}$$

given by

$$\sum_{a=0}^{\infty} b_a \ \mapsto \ll \sum_{a=0}^{\infty} r^a b_a / E_{p-1}^a \gg, \tag{6.60}$$

where the right hand side stands for the p-adic modular form whose value on a test object $(E, \omega, \beta_N, Y)_{/A}$ (where $Y \cdot E_{p-1}(\underline{A}, \omega) = r$) is:

$$\ll \sum_{a=0}^{\infty} b_a / E_{p-1}^a \gg (E, \omega, \beta_N, Y) = \sum_{a=0}^{\infty} Y^a b_a(E, \omega, \beta_N). \tag{6.61}$$

PROOF. One easily sees that ψ is a well-defined continuous homomorphism of B-modules (in particular, the sum on the right hand side converges). Note that $\mathbb{M}(B,k,\mu_N;r)$ is the p-adic completion of $\left(\oplus_{j=0}^{\infty}\mathcal{M}(B,k+j(p-1),\mu_N)\right)/(E_{p-1}-r)$.

Consider the following diagram:

$$\begin{array}{ccc} A^{\mathrm{rigid}}(B,k,\mu_N) & \xrightarrow{\quad\psi\quad} & \mathbb{M}(B,k,\mu_N;r) \\ \uparrow{\scriptstyle\xi} & & \uparrow{\scriptstyle\phi} \\ \overset{j}{\underset{a=0}{\oplus}} A(B,k+a(p-1),\mu_N) & \xrightarrow{\quad\eta\quad} & \mathcal{M}(B,k+j(p-1),\mu_N) \end{array} \tag{6.62}$$

In this diagram:

$$\eta(\sum_{a=0}^{j} b_a) = \sum_{a=0}^{j} b_a \cdot E_{p-1}^{j-a} \quad \text{(an isomorphism);} \tag{6.63}$$

$$\xi(\sum_{a=0}^{j} b_a) = \sum_{a=0}^{j} b_a \quad \text{(the natural inclusion);} \tag{6.64}$$

$$\psi(\sum_{a=0}^{\infty} b_a)(E,\omega,\beta_N,Y) = \sum_{a=0}^{\infty} Y^a b_a(E,\omega,\beta_N); \tag{6.65}$$

$$\phi(f)(E,\omega,\beta_N,Y) = Y^j \cdot f(E,\omega,\beta_N). \tag{6.66}$$

The diagram commutes:

$$\phi(\eta(\sum_{a=0}^{j} b_a))(E,\omega,\beta_N,Y) = \phi(\sum_{a=0}^{j} b_a \cdot E_{p-1}^{j-a})(E,\omega,\beta_N,Y) \tag{6.67}$$

$$= Y^j(\sum_{a=0}^{j} b_a E_{p-1}^{j-a})(E,\omega,\beta_N) \tag{6.68}$$

$$= \sum_{a=0}^{j} Y^a b_a(E,\omega,\beta_N) \tag{6.69}$$

$$= \psi(\xi(\sum_{a=0}^{j} b_a))(E,\omega,\beta_N,Y). \tag{6.70}$$

Now, every element x of $\mathbb{M}(B,k,\mu_N;r)$ can be written as

$$\sum_{j=0}^{\infty} s_j, \quad s_j \in M(B,k+j(p-1),\mu_N), \quad s_j \longrightarrow 0 \ \ p\text{-adically.} \tag{6.71}$$

Note that when we say $\sum_{j=0}^{\infty} s_j \in \mathbb{M}(B,k,\mu_N;r)$, we really mean that as a function whose values on a test object are given by $(\sum s_j)(E,\omega,\beta_N,Y) := \sum Y^j s_j(E,\omega,\beta_N)$. Hence, this element really is $\psi(\sum s_j)$ *if* the sum belongs to $A^{\mathrm{rigid}}(B,k,\mu_N)$!

Improving on Equation (6.71), one can write x as:

$$\sum_{j=0}^{\infty} s_j, \quad s_j \in A(B, k+j(p-1), \mu_N). \tag{6.72}$$

Moreover, still $s_j \longrightarrow 0$ p-adically. Indeed, if p^N divides $s_j = a_j + b_j$, where $a_j \in A(B, k+j(p-1), \mu_N)$ and $b_j \in E_{p-1} \cdot \mathcal{M}(B, k+(j-1)(p-1), \mu_N)$ then $p^N | b_j$. Thus, collecting terms backwards does not destroy convergence! $s_j \longrightarrow 0$ p-adically, and thus $\sum_{j=0}^{\infty} s_j \in A^{\mathrm{rigid}}(B, k, \mu_N)$. This shows that ψ is surjective.

To prove the injectivity of ψ, assume that $\psi(\sum_{a=0}^{\infty} b_a) = 0$. Fix an integer N_0. Then for some N_1 we have $a > N_1 \Rightarrow p_{N_0} | b_a$. Thus

$$\psi(\sum_{a=0}^{N_1} b_a) \equiv 0 \pmod{p^{N_0}}. \tag{6.73}$$

However, $\psi(\sum_{a=0}^{N_1} b_a) = \phi(\eta(\sum_{a=0}^{N_1} b_a)) = \phi(\sum_{a=0}^{N_1} b_a E_{p-1}^{N_1-a})$. Therefore,

$$\phi(\sum_{a=0}^{N_1} b_a E_{p-1}^{N_1-a}) \equiv 0 \pmod{p^{N_0}}. \tag{6.74}$$

But,

$$\mathcal{M}(B, k+N_1(p-1), \mu_N)/(p^{N_0}) = \mathcal{M}(B/(p^{N_0}), k+N_1(p-1), \mu_N) \hookrightarrow \mathbb{M}(B/(p^{N_0}), k, \mu_N; r) = \mathbb{M}(B, k, \mu_N; r)/(p^{N_0}). \tag{6.75}$$

Thus $\sum_{a=0}^{N_1} b_a \equiv 0 \pmod{p^{N_0}}$, and hence $\sum_{a=0}^{\infty} b_a \equiv 0 \pmod{p^{N_0}}$ for every N_0. This implies $\sum_{a=0}^{\infty} b_a = 0$. □

COROLLARY 6.11. *Let $r_2 = rr_1$. Then the natural map*

$$\mathbb{M}(B, k, N; r_2) \longrightarrow \mathbb{M}(B, k, N; r_1), \tag{6.76}$$

(given by $f \mapsto f'$ and $f'(E, \omega, \beta_N, Y) = f(E, \omega, \beta_N, rY)$) is given by

$$\sum_{a=0}^{\infty} b_a \mapsto \sum_{a=0}^{\infty} r^a b_a \tag{6.77}$$

and is thus injective.

6.6. Properties of q-expansions of p-adic modular forms.

PROPOSITION 6.12. *Let $b \in B$ be an element dividing a positive power of p. Let $f \in \mathbb{M}(B, k, \mu_N; 1)$. The followings assertions are equivalent:*

1. *$f \in b \cdot \mathbb{M}(B, k, \mu_N; 1)$.*
2. *The q-expansion of f lies in $b \cdot B[[q]]$.*

PROOF. Since the q-expansion is B-linear, the implication $1 \implies 2$ is clear. Now, let us prove $2 \implies 1$. Note that

$$\mathbb{M}(B/bB, k, \mu_N; 1) = A^{\mathrm{rigid}}(B/bB, k, \mu_N) = A^{\mathrm{rigid}}(B, k, \mu_N)/bA^{\mathrm{rigid}}(B, k, \mu_N). \tag{6.78}$$

Thus, replacing B by B/bB, we may assume that $b = 0$ and that p is nilpotent in B. We aim to prove that the q-expansion map is injective.

For p nilpotent, f is a *finite* sum $\sum_{a=0}^{d} b_a$, $b_a \in A(B, k+a(p-1), \mu_N)$, and its q-expansion is

$$f(\mathbf{Tate}(q), \omega_{can}, \beta_{can}, E_{p-1}^{-1}) = \sum_{a=0}^{d} E_{p-1}^{-a} \cdot (q - \text{expansion of } b_a) \tag{6.79}$$

$$= \sum_{a=0}^{d} \frac{q - \text{expansion of } (E_{p-1}^{d-a} \cdot b_a)}{E_{p-1}^{d}}, \tag{6.80}$$

so the q-expansion of $\sum_{a=0}^{d} E_{p-1}^{d-a} \cdot b_a$ is zero. By the q-expansion principal for classical modular forms $\sum_{a=0}^{d} E_{p-1}^{d-a} \cdot b_a$ is zero. This implies that each b_a is zero, because $(b_0, \ldots, b_d) \in \sum_{a=0}^{M} A(B, k+a(p-1), \mu_N) \cong \mathcal{M}(B, k+d(p-1), \mu_N)$. □

Corollary 6.13. *The q-expansion map on p-adic modular forms is injective.*

Theorem 6.14. *Let $f(q) \in B[[q]]$ be a power series. The following assertions are equivalent:*

1. *$f(q)$ is the q-expansion of an element $f \in \mathbb{M}(B, k, \mu_N; 1)$.*
2. *For all n, there exists a positive integer $M(n)$, $M(n) \equiv 0 \mod p^{n-1}$, and a classical modular form $g_n \in \mathcal{M}(B, k+M(n)(p-1), \mu_N)$ such that the q-expansion $g_n(q) \equiv f(q) \mod p^n$.*

Proof. First, let us show that 2 $\implies$ 1. Writing $E_{p-1}(q) = 1 + px$, we see that $E_{p-1}^{p^{n-1}} \equiv 1 \mod p^n$. Now, multiplication of g_n by $E_{p-1}^{p^{n-1}}$ changes the weight by $(p-1)p^{n-1}$, so we can assume $M(n)$ is increasing. Let $\Delta(n) \equiv M(n+1) - M(n)$, so

$$g_{n+1} - g_n \cdot E_{p-1}^{\Delta(n)} \in p^n \cdot \mathcal{M}(B, k+(p-1)M(n+1), \mu_N), \tag{6.81}$$

(since $\Delta(n) \equiv 0 \mod p^{n-1}$).

Hence, $g_0 + \sum_{a=0}^{\infty} (g_{a+1} - g_a \cdot E_{p-1}^{\Delta(a)}) \in \mathbb{M}(B, k, \mu_N; 1)$. Modulo p^n this sum is

$$g_0 + (g_1 - g_0 E_{p-1}^{\Delta(0)}) + \cdots + (g_n - g_{n-1} \cdot E_{p-1}^{\Delta(n-1)}). \tag{6.82}$$

But $E_{p-1} = 1$ in $\mathbb{M}(B, k, \mu_N; 1)$, so the telescopic sum is equal to g_n. Hence, the q-expansion is $\lim g_n(q) = f(q)$.

The implication 1 $\implies$ 2 can be proved as follows: Let $f \in \mathbb{M}(B, k, \mu_N; 1)$. Then

$$f = \psi(\sum_{a=0}^{\infty} b_a), \quad b_a \in A(B, k+a(p-1), \mu_N). \tag{6.83}$$

Consider $c_n = \psi(\sum_{a=0}^{n} b_a) = \phi(\eta(\sum_{a=0}^{n} b_a)) \in \mathcal{M}(B, k+n(p-1), \mu_N)$. Take $M(n)$ to be suitably increasing powers of p and $g_n = c_{M(n)}$. □

Corollary 6.15. *Serre's p-adic modular forms of weight $k \in \mathbb{Z}$ are the same as p-adic modular forms à la Katz of growth condition 1: $\mathbb{M}(B, k, \mu_N; 1)$.*

7. The Ring of Divided Congruences

In this section we follow Katz [**59**]. Let $p \geq 5$ be a prime number and let k be a perfect field of characteristic p. We fix the following notation

$W(k)$	the ring of infinite Witt vectors over k.
$W_m(k)$	the ring of Witt vectors of length m over k (equal to $W(k)/p^m W(k)$).
N	an integer ≥ 3 and prime to p (auxiliary level).
ζ	a fixed root of unity of order N in k (or in $W(k)$ via the Teichmüller lift).
$\mathfrak{M}(W(k), N)$	the fine moduli scheme of elliptic curves over $W(k)$ algebras with symplectic level N structure ($(\mathbb{Z}/n\mathbb{Z})^2 \cong E[N]$ s.t. the symplectic pairing given by $< (1,0),(0,1) >= \zeta$ corresponds to the Weil pairing).
$\mathfrak{M}^*(W(k), N)$	the canonical compactification of $\mathfrak{M}(W(k), N)$ obtained by adding cusps.
M_m^0	is $\mathfrak{M}(W_m(k), N) = \mathfrak{M}(W(k), N) \otimes_{W(k)} W_m(k)$.
S_m^0	is the open subscheme of M_m^0 where the Hasse invariant is invertible.
M_m	is $\mathfrak{M}^*(W_m(k), N) = \mathfrak{M}^*(W(k), N) \otimes_{W(k)} W_m(k)$.
S_m	is the open subscheme of M_m where the Hasse invariant is invertible.

We remark that the schemes S_m and S_m^0 are affine and that the structural morphisms $S_m \longrightarrow W_m(k)$ and $S_m^0 \longrightarrow W_m(k)$ are smooth with an irreducible special fiber. We have the compatibilities:

$$S_m = S_{m+1} \otimes_{W_{m+1}(k)} W_m(k), \quad S_m^0 = S_{m+1}^0 \otimes_{W_{m+1}(k)} W_m(k). \tag{7.1}$$

We let

$$T_{m,n} \longrightarrow S_m \tag{7.2}$$

be the étale covering of μ_{p^n}-level: $\beta_{p^n} : \mu_{p^n} \hookrightarrow E$. That is, $T_{m,n}$ represents the moduli functor of elliptic curves over $W_m(k)$-algebras with symplectic level N structure and μ_{p^n} structure. The covering $T_{m,n} \longrightarrow S_m$ is thus Galois with Galois group $(\mathbb{Z}/p^n\mathbb{Z})^\times$. It follows that $T_{m,n}$ is an affine scheme and that the morphism $T_{m,n} \longrightarrow W_m(k)$ is smooth. We have the compatibility

$$T_{m,n} = T_{m+1,n} \otimes_{W_{m+1}(k)} W_m(k). \tag{7.3}$$

The schemes $T_{1,n}$ appeared before in Section 6. We offer a panoramic view:

$$\begin{array}{ccccccc}
\vdots & & \vdots & & \vdots & & \\
\downarrow & & \downarrow & & \downarrow & & \\
T_{1,3} & \longrightarrow & T_{2,3} & \longrightarrow & T_{3,3} & \longrightarrow & \dots \\
{\scriptstyle \mathbb{Z}/p\mathbb{Z}}\downarrow & & \downarrow & & \downarrow & & \\
T_{1,2} & \longrightarrow & T_{2,2} & \longrightarrow & T_{3,2} & \longrightarrow & \dots \\
{\scriptstyle \mathbb{Z}/p\mathbb{Z}}\downarrow & & \downarrow & & \downarrow & & \\
T_{1,1} & \longrightarrow & T_{2,1} & \longrightarrow & T_{3,1} & \longrightarrow & \dots \\
{\scriptstyle (\mathbb{Z}/p\mathbb{Z})^\times}\downarrow & & \downarrow & & \downarrow & & \\
S_1 & \longrightarrow & S_2 & \longrightarrow & S_3 & \longrightarrow & \dots
\end{array} \tag{7.4}$$

The horizontal arrows are all closed immersions. The vertical arrows are all étale Galois coverings. Let $V_{m,n}$ be the ring of regular functions on $T_{m,n}$. We get the dual diagram, with horizontal arrows being surjective and vertical arrows being inclusions.

$$\begin{array}{ccccccc}
\vdots & & \vdots & & \vdots & & \\
\uparrow & & \uparrow & & \uparrow & & \\
V_{1,3} & \longleftarrow & V_{2,3} & \longleftarrow & V_{3,3} & \longleftarrow & \dots \\
\uparrow & & \uparrow & & \uparrow & & \\
V_{1,2} & \longleftarrow & V_{2,2} & \longleftarrow & V_{3,2} & \longleftarrow & \dots \\
\uparrow & & \uparrow & & \uparrow & & \\
V_{1,1} & \longleftarrow & V_{2,1} & \longleftarrow & V_{3,1} & \longleftarrow & \dots \\
\uparrow & & \uparrow & & \uparrow & & \\
V_{1,0} & \longleftarrow & V_{2,0} & \longleftarrow & V_{3,0} & \longleftarrow & \dots
\end{array} \tag{7.5}$$

Note that

$$V_{m+1,n}/p^m V_{m+1,n} \cong V_{m,n}. \tag{7.6}$$

We let

$$T_{m,\infty} = \varprojlim_n T_{m,n}, \quad T_{\infty,\infty} = \varinjlim_m T_{m,\infty}. \tag{7.7}$$

We remark that the scheme $T_{\infty,\infty}$ represents the functor of $W(k)$- algebras

$$R \mapsto \text{Iso. classes of } (E, \beta_{p^\infty} : \widehat{\mathbb{G}_m} \xrightarrow{\cong} \widehat{E}). \tag{7.8}$$

Let

$$V_{m,\infty} = \varinjlim_n V_{m,n}, \quad V_{\infty,\infty} = \varprojlim_m V_{m,\infty}. \tag{7.9}$$

Note that

$$V_{m+1,\infty}/p^m V_{m+1,\infty} \cong V_{m,\infty}. \tag{7.10}$$

The rings $V_{m,n}$ are smooth W_m algebras and via the standard cusp we have

$$V_{m,n} \subset W_m[[q]]. \tag{7.11}$$

LEMMA 7.1. *The cokernel $W_m[[q]]/V_{m,n}$ is a flat W_m-module.*

PROOF. Since W_m is a local ring, the assertion amounts to

$$p \cdot f(q) \in V_{m,n} \Rightarrow f(q) \in V_{m,n}. \tag{7.12}$$

Let $g \in V_{m,n}$ with a q-expansion $p \cdot f(q)$. Now

$$V_{m,n} \otimes_{W_m} W_1 = V_{m,n} \otimes_{W_m} k = V_{1,n} \hookrightarrow k[[q]]. \tag{7.13}$$

Thus g is zero modulo $pV_{m,n}$. □

COROLLARY 7.2. *For $m \leq \infty$ we have an injection*

$$V_{m,\infty} \hookrightarrow W_m[[q]] \tag{7.14}$$

with a W_m flat cokernel.

Consider the group scheme $\mu_{p^n}{}_{/A} = \operatorname{Spec} A[x]/(x^{p^n}-1)$. Let $B = A[x]/(x^{p^n}-1)$. Then

$$\Omega_{\mu_{p^n}/\operatorname{Spec}(A)} = \Omega_{B/A} \tag{7.15}$$

is the B-module generated by dx with the relation $d(x^{p^n}-1)=0$. That is, $\Omega_{B/A}$ is the B-module

$$B[x]/(x^{p^n}-1, p^n x^{p^n-1}). \tag{7.16}$$

It is free over B iff $p^n = 0$ in B.

We conclude that the schemes $T_{m,n}$ with $m \leq n$ are special in that that the level structure

$$\beta_{p^n} : \mu_{p^n}{}_{/R} \longrightarrow A_{/R}, \qquad R \in \mathbf{W_m} - \mathbf{Alg}, \tag{7.17}$$

induces an isomorphism

$$\mathfrak{t}^*_{\mu_{p^n}/R} \cong \mathfrak{t}^*_{A/R}. \tag{7.18}$$

In particular, *there exists a canonical modular form of weight* 1, $a(1)$ *on* $T_{m,n}$ *for* $m \leq n$ *corresponding to the image of the canonical generator* dx/x *of* $\mathfrak{t}^*_{\mu_{p^n}/R}$.

EXERCISE 7.3. Prove that under the maps

$$T_{m,n+1} \longrightarrow T_{m,n} \quad m \leq n \tag{7.19}$$

and

$$T_{m,n} \longrightarrow T_{m+1,n} \quad m+1 \leq n \tag{7.20}$$

the modular forms $a(1)$ agree. Conclude that there exists a modular form of weight 1 on the schemes $T_{m,\infty}$ and $T_{\infty,\infty}$.

We let

$$a(i) := a(1)^i, \tag{7.21}$$

denote the modular form of weight i on the scheme $T_{m,n}$ with $m \leq n$ (including $m = \infty$).

For every m let

$$R_m = \bigoplus_{k \geq 0} \Gamma(M_m, \underline{\omega}^k) \tag{7.22}$$

be the graded ring of classical modular forms on the scheme M_m ($\omega = \pi_* \mathfrak{t}^*_{\mathcal{E}^U/M_m}$). Let

$$R_\infty = \bigoplus_{k \geq 0} \Gamma(M, \underline{\omega}^k), \tag{7.23}$$

be the graded ring of classical modular forms on the scheme M.

Now, fix m and define a homomorphism

$$r_m : R_m \longrightarrow V_{m,m} \hookrightarrow V_{m,\infty}, \tag{7.24}$$

by

$$r_m(\sum f_i) = \sum f_i/a(i), \quad f_i \in T(M_m, \underline{\omega}^i). \tag{7.25}$$

We define

$$r_\infty : R_\infty \longrightarrow V_{\infty,\infty}, \tag{7.26}$$

by the composition

$$R_\infty \longrightarrow \varprojlim_m R_m \xrightarrow{\varprojlim r_m} \varprojlim_m V_{m,m} = V_{\infty,\infty} \ . \tag{7.27}$$

Since the q-expansion of $a(1)$ is one at the standard cusp, the q expansion of $r_m(\sum f_i)$ is just $\sum f_i(q)$, where $f_i(q)$ is the q-expansion of f_i and it belongs to $W_m[[q]]$.

Let

$$I_{m,m_1} \lhd R_m \tag{7.28}$$

be the ideal of R_m consisting of all sums $\sum f_i$ such that $\sum f_i(q) \equiv 0 \pmod{p^{m_1}}$ at the standard cusp.

LEMMA 7.4. *Let $\sum f_i \in I_{m,m_1}$, then the q expansion $\sum f_i(q)$ with respect to* any *cusp satisfies $\sum f_i(q) \equiv 0 \mod p^{m_1}$.*

PROOF. We are given that $r_m(p^{m-m_1}\sum f_i)$, which is a *function* on the irreducible scheme $T_{m,m}$, is expressed in the local ring a certain point (i.e., the one belonging to the standard cusp) by zero. Therefore, the function $r_m(p^{m-m_1}\sum f_i)$ *is* zero, and hence is zero in every local ring on $T_{m,m}$. In particular, it is zero in the local ring of any other cusp. That is, all its q-expansions are zero. □

COROLLARY 7.5. *Let $\sum f_i \in I_{m,m_1}$. Then for all $a \in (\mathbb{Z}/p^m\mathbb{Z})^\times$*

$$\sum a^i f_i(q) \cong 0 \mod p^{m_1}. \tag{7.29}$$

PROOF. The Galois action of $(\mathbb{Z}/p^m\mathbb{Z})^\times$ on $V_{m,m}$ is best described on points: The effect of $a \in (\mathbb{Z}/p^m/Z)^\times$ is $(E, \beta_{p^n}) \mapsto (E, \beta_{p^n} \circ a)$. As in the proof of Theorem 5.2 one verifies that r_m is equivariant with respect to this action when we let a act on a modular form of weight i by a^i. □

Let $\sum f_i \in I_{m,m_1}$ $(m_1 \le m)$. Then $p^{m-m_1} r_m(\sum f_i)$ is zero in V_{m,m_1}. Since $W_m[[q]]/V_{m,m_1}$ is a flat W_m-module, there exists an $h \in V_{m,m_1}$ such that

$$r_m(\sum f_i) = p^{m_1} h. \tag{7.30}$$

The function h is unique modulo p^{m-m_1} and we obtain a well defined homomorphism

$$\gamma_{m_1} : I_{m,m_1} \longrightarrow V_{m-m_1,m} \subset V_{m-m_1,\infty}. \tag{7.31}$$

We write symbolically,

$$\gamma_{m_1} = \frac{1}{p^{m_1}} r_m. \tag{7.32}$$

Passing to the limit, we get a homomorphism

$$\gamma_{m_1} : I_{\infty,m_1} \longrightarrow V_{\infty,\infty}. \tag{7.33}$$

If $\sum f_i \in I_{\infty,m_1}$ then $\sum f_i \in R_\infty = \oplus_k \Gamma(M, \underline{\omega}^k)$ and $\sum f_i(q) \equiv 0 \pmod{p^{m_1}}$; furthermore,

$$(\gamma_{m_1}(\sum f_i))(q) = \frac{1}{p^{m_1}} \sum f_i(q). \tag{7.34}$$

We have the following diagram

$$\begin{array}{ccccccccc} R_\infty & \longrightarrow & \frac{1}{p} I_{\infty,1} & \longrightarrow & \frac{1}{p^2} I_{\infty,2} & \longrightarrow & \dots & \longrightarrow & D \subset R_\infty \left[\frac{1}{p}\right] \\ \downarrow r_\infty & & \downarrow p\cdot\gamma_1 & & \downarrow p^2\cdot\gamma_2 & & \dots & & \downarrow \beta \\ V_{\infty,\infty} & \longrightarrow & V_{\infty,\infty} & \longrightarrow & V_{\infty,\infty} & \longrightarrow & \dots & \longrightarrow & V_{\infty,\infty} \end{array} \tag{7.35}$$

(Horizontal arrows being the natural inclusions). The $\mathbb{Z}_p$-algebra D is precisely the elements in $R_\infty \left[\frac{1}{p}\right]$ with p-integral q-expansion. It is called *the ring of divided congruences*.

THEOREM 7.6. *For every m, the map*

$$\beta(m) : D/p^m D \longrightarrow V_{\infty,\infty}/p^m V_{\infty,\infty} = V_{m,\infty}, \tag{7.36}$$

induced from β, is an isomorphism.

PROOF.

• $\beta(m)$ is injective.

Let $\sum f_i \in p^{-n} I_{\infty,n}$ be in the kernel of $\beta(m)$. Then $\sum f_i(q) \equiv 0 \pmod{p^m}$. Therefore, $\sum f_i \in p^{-n} I_{\infty,n+m} = p^m (p^{-n-m} I_{\infty,n+m}) \subset p^m D$.

• $\beta(m)$ is surjective.

We claim that if $\beta(1) : D/pD \longrightarrow V_{1,\infty}$ is surjective then $\beta(m) : D/p^m D \longrightarrow V_{m,\infty}$ is surjective. Indeed, if $\beta(1) : D/pD \longrightarrow V_{1,\infty}$ is surjective and $x \in V_{m,\infty}$, there exists an element $g \in D$ such that $\beta(m)(g) \equiv x \pmod{pV_{m,\infty}}$. Therefore $\beta(m)(g) - x = px_1$ for some $x_1 \in V_{m,\infty}$, etc.. The process stops because p is nilpotent in $V_{m,\infty}$. It remains to show that

• $\beta(1)$ is surjective.

We know that the composition

$$R_\infty \longrightarrow V_{1,\infty} \longrightarrow V_{1,1} \tag{7.37}$$

is surjective (see the proof of Theorem 5.2). We shall abuse notation and denote this map by $\beta(1)$ as well.

Artin-Schreier theory says the following: Let A be a ring of characteristic p (e.g. $A = V_{1,n}$) and $B \supset A$ a finite étale A algebra with Galois group $\mathbb{Z}/p\mathbb{Z}$ (e.g. $B = V_{1,n+1}$), then there exists $b \in B$ such that: (i) $b^p - b \in A$; (ii) $\ell \in \mathbb{Z}/p\mathbb{Z}$ acts by $b \mapsto b + \ell$. Such an element b is unique up to addition of elements from A. Note that any element $b \in B$ such that $b \mapsto b + 1$ under some element of $\mathbb{Z}/p\mathbb{Z}$ generated B as an A-module. We call such an element an *Artin-Schreier generator* for B over A.

We next note that the group $\mathbb{Z}_p^\times$ acts on D. First, $\mathbb{Z}_p^\times$ acts on $R_\infty[1/p]$ by

$$[a](\sum f_i) = \sum a^i f_i. \tag{7.38}$$

By Corollary 7.5 the action preserves D.

LEMMA 7.7. (Key Lemma) *Let $n \geq 1$. There exists an element $d_n \in D$ such that for every $k \geq 0$*

$$[1+p^{n+k}](d_n) \equiv d_n + p^k E_{p-1} \pmod{p^{k+1}D}. \tag{7.39}$$

We first explain how the Theorem follows from the Key Lemma. We know that

$$\beta(1) : R_\infty \twoheadrightarrow V_{1,1}. \tag{7.40}$$

It follows that for $n \geq 1$ we have

$$V_{1,n+1} = V_{1,n}[\beta(1)(d_n)], \tag{7.41}$$

where d_n is the element provided by the Key Lemma. Indeed,

$$[1+p^{n+1}](d_n) \equiv d_n + pE_{p-1} \pmod{p^2 D} \tag{7.42}$$

and

$$d_n + pE_{p-1} \equiv d_n \pmod{pD}, \tag{7.43}$$

whence $\beta(1)(d_n) \in V_{1,n+1}$ (because $\mathrm{Aut}(V_{1,n+1}/V_{1,n}) = (1+p^n\mathbb{Z}_p)/(1+p^{n+1}\mathbb{Z}_p)$). Similarly,

$$[1+p^n](d_n) \equiv d_n + E_{p-1} \equiv d_n + 1 \pmod{pD}. \tag{7.44}$$

Thus, d_n is an Artin-Schreier generator for $V_{1,n+1}$ over $V_{1,n}$ and (7.41) follows. Therefore, the map

$$\beta(1) : R_\infty \longrightarrow V_{1,n} \tag{7.45}$$

is surjective for every n and that implies that

$$\beta(1) : R_\infty \longrightarrow V_{1,\infty} \tag{7.46}$$

is surjective because R_∞ is p-adically complete.

Proof of Key Lemma. The proof is by induction on n. For $n = 1$ take

$$d_1 = \frac{1 - E_{p-1}}{p}. \tag{7.47}$$

We calculate that for a suitable integer c

$$[1+p^{1+k}](d_1) = [1+p^{1+k}]\left(\frac{1-E_{p-1}}{p}\right) \tag{7.48}$$

$$= \frac{1 - (1+p^{1+k})^{p-1}E_{p-1}}{p} \tag{7.49}$$

$$= \frac{1-E_{p-1}}{p} - \frac{(p-1)p^{1+k}E_{p-1}}{p} - \frac{p^{2+2k}}{p} \cdot c \cdot E_{p-1} \tag{7.50}$$

$$= d_1 + p^k E_{p-1} \pmod{p^{k+1}D}. \tag{7.51}$$

Suppose that $d_1, \ldots, d_n$ were constructed. As noted, for $\ell \leq n$

$$V_{1,\ell+1} = V_{1,\ell}[\beta(1)d_\ell], \;\; (\beta(1)d_\ell) - (\beta(1)d_\ell)^p \in V_{1,\ell}. \tag{7.52}$$

Thus,

$$V_{1,n} = \beta(1)(R_\infty[d_1, \ldots, d_{n-1}]), \tag{7.53}$$

and

$$(\beta(1)d_n) - (\beta(1)d_n)^p = \beta(1)c_n \tag{7.54}$$

for some $c_n \in R_\infty[d_1, \ldots, d_{n-1}]$. Hence, $d_n - d_n^p - c_n \in \mathrm{Ker}(\beta(1)) = pD$. We let

$$d_{n+1} = \frac{d_n - d_n^p - c_n}{p}. \tag{7.55}$$

It is an element of D. We check the Galois action. First:

$$[1 + p^{(n+1)+k}](d_{n+1}) = \frac{[1 + p^{(n+1)+k}](d_n) - \left([1 + p^{(n+1)+k}](d_n)\right)^p - [1 + p^{(n+1)+k}](c_n)}{p}. \tag{7.56}$$

On the other hand, by induction,

- $[1 + p^{n+(1+k)}](d_n) \equiv d_n + p^{k+1}E_{p-1} \pmod{p^{k+2}D}$.
- $\left([1 + p^{n+(1+k)}](d_n)\right)^p \equiv d_n^p \pmod{p^{k+2}D}$.
- $[1 + p^{n+(1+k)}](c_n) \equiv c_n \pmod{p^{k+2}D}$ (true for any $c \in R_\infty[d_1, \ldots, d_n]$).

Put together this yields

$$[1 + p^{(n+1)+k}](d_{n+1}) \equiv \frac{d_n - d_n^p - c_n + p^{k+1}E_{p-1}}{p} \tag{7.57}$$

$$\equiv d_{n+1} + p^k E_{p-1} \pmod{p^{k+1}D}. \tag{7.58}$$

□

Corollary 7.8. *The ring of divided congruences D is p-adically dense in the space of Katz functions $V_{\infty,\infty}$.*

CHAPTER 5

p-adic Hilbert Modular Forms

In this chapter we follow the conceptual frame work laid in Chapter 4. We shall not be able to present a theory as complete as for elliptic curves simply because the theory is not yet worked out!

The motivation for studying p-adic Hilbert modular forms includes the same reasons given in Chapter 4. Namely, they include p-adic interpolation of L-functions associated to totally real fields (see Cassou-Noguès [**7**] and Deligne-Ribet [**23**]), and deformations of Galois representations of a totally real field.

When developing the theory of p-adic modular forms one must abandon the hope of using specific knowledge of the ring of modular forms over $\mathbb{C}$ and using "down to earth arguments" as, say, in Chapter 4, Section 2. That distinguishes the case of $g > 1$ from the elliptic case. On a conceptual level this is expected: while for $g = 1$ there is "one modular curve" (i.e. up to correspondences; or $\varprojlim \Gamma(n)\backslash\mathfrak{H}$), for $g > 1$ there is "one modular variety" (in the same sense) for every totally real field of degree g over $\mathbb{Q}$.

1. Algebraic Hilbert Modular Forms

Let $N \geq 4$ be an integer. Let

$$\mathfrak{M}(\mu_N) \longrightarrow \mathrm{Spec}(\mathcal{O}_L[d_L^{-1}]) \tag{1.1}$$

be the fine moduli scheme parameterizing abelian varieties with real multiplication and μ_N level structure $\underline{A} = (A, \iota, \beta_N)$ satisfying condition (**R**) (it exists in fact over $\mathbb{Z}[d_L^{-1}]$). See Chapter 4, Section 5.

For any scheme $f : B \longrightarrow \mathrm{Spec}(\mathcal{O}_L[d_L^{-1}])$ let

$$\mathfrak{M}(B, \mu_N) = \mathfrak{M}(\mu_N) \times_{\mathrm{Spec}(\mathcal{O}_L[d_L^{-1}])} B \longrightarrow B. \tag{1.2}$$

It is the fine moduli space for abelian varieties with real multiplication and μ_N level structure over base schemes $S \in \mathbf{Sch}_{/\mathbf{B}}$. If $B = \mathrm{Spec}(R)$ we just write $\mathfrak{M}(R, \mu_N)$. Let

$$\pi : \underline{A}^U \longrightarrow \mathfrak{M}(\mu_N) \tag{1.3}$$

be the universal object and let $\underline{A}^B$ be the induced universal object over $\mathfrak{M}(B, \mu_N)$. We let

$$\underline{\omega} = \mathfrak{t}^*_{\underline{A}^U/\mathfrak{M}(\mu_N)} \tag{1.4}$$

be the sheaf over $\mathfrak{M}(\mu_N)$ of relative cotangent spaces at the origin. Similarly, for $f : B \longrightarrow \mathrm{Spec}(\mathcal{O}_L[d_L^{-1}])$ we let

$$\underline{\omega}_B = \mathfrak{t}^*_{\underline{A}^B/B} = f^*\underline{\omega}. \tag{1.5}$$

When there is no danger of confusion we write $\underline{\omega}$ for $\underline{\omega}_B$.

The sheaf $\underline{\omega}_B$ is a locally free $\mathcal{O}_B$-module of rank g and is a locally free $\mathcal{O}_L \otimes \mathcal{O}_B$-module of rank one.

Let B be a $\mathcal{O}_L[d_L^{-1}]$-algebra. In order to define weights for Hilbert modular forms we introduce the following functor $\mathbb{T}_B$. It is a functor from B-algebras to abelian groups given by

$$R \mapsto \mathbb{T}_B(R) = (\mathcal{O}_L \otimes_{\mathbb{Z}} R)^{\times}, \qquad R \in \mathbf{B\text{-}Alg}. \tag{1.6}$$

Then $\mathbb{T}_B$ is a group scheme over $\mathrm{Spec}(B)$, which is in fact a torus. [1] That is, for every geometric point $\mathrm{Spec}(k) \longrightarrow \mathrm{Spec}(B)$ the scheme $\mathbb{T} \times_{\mathrm{Spec}(B)} \mathrm{Spec}(k)$, which represents the functor

$$R \mapsto (\mathcal{O}_L \otimes R)^{\times} \cong \overset{g}{\underset{i=1}{\oplus}} R^{\times}, \qquad R \in \mathbf{k\text{-}Alg}, \tag{1.7}$$

is isomorphic to ${\mathbb{G}_m}^g_{/k}$.

Let $\tilde{L}$ be a Galois closure of L. If we take B to be an $\mathcal{O}_{\tilde{L}}[d_L^{-1}]$-algebra, where $\mathcal{O}_{\tilde{L}}$ is the ring of integers of $\tilde{L}$, then $\mathbb{T}_B$ is a *split* torus: if we enumerate the embeddings

$$\sigma_1, \cdots, \sigma_g : L \hookrightarrow \tilde{L}, \tag{1.8}$$

then

$$\mathcal{O}_L \otimes B = \overset{g}{\underset{i=1}{\oplus}} B, \tag{1.9}$$

by the map induced from

$$a \otimes 1 \mapsto (\sigma_1(a), \ldots, \sigma_g(a)), \tag{1.10}$$

and we get a canonical isomorphism

$$\mathbb{T}_B = {\mathbb{G}_m}^g_{/B}. \tag{1.11}$$

Assume henceforth that B is an $\mathcal{O}_{\tilde{L}}[d_L^{-1}]$-algebra. The character group $X(\mathbb{T}_B)$ is given by the free abelian group (written multiplicatively)

$$X(\mathbb{T}_B) =< \chi_1, \ldots, \chi_g > \tag{1.12}$$

on the characters $\chi_1, \ldots, \chi_g$, where χ_i is the projection of $\mathbb{T}_B$ on the $i-th$ component

$$\mathbb{T}_B = {\mathbb{G}_m}^g_{/B} \longrightarrow \mathbb{G}_{m/B}. \tag{1.13}$$

We are ready to define modular forms. Over $\mathfrak{M}(B, \mu_N)$ the vector bundle $\underline{\omega}$ decomposes as a direct sum of line bundles

$$\underline{\omega} = \underline{\omega}(\chi_1) \oplus \cdots \oplus \underline{\omega}(\chi_g). \tag{1.14}$$

This decomposition comes from Equation (1.9), $\mathcal{O}_L \otimes B = \oplus_{i=1}^g B$, where we also denote the i-th projection from $\mathcal{O}_L \otimes B$ to B by

$$\chi_i : \mathcal{O}_L \otimes B \longrightarrow B. \tag{1.15}$$

[1] In fact $\mathbb{T}_B = (\mathbf{Res}_{\mathcal{O}_L/\mathbb{Z}} \mathbb{G}_m) \times_{\mathrm{Spec}(\mathbb{Z})} \mathrm{Spec}(B)$.

Note that $\chi_i(a \otimes 1) = \sigma_i(a)$ as an element of B. Thus, $\mathcal{O}_L$ acts on $\underline{\omega}(\chi_i)$ via the homomorphism $\chi_i : \mathcal{O}_L \otimes B \longrightarrow B$. We define for $\chi = \chi_1^{a_1} \cdots \chi_g^{a_g}$

$$\underline{\omega}(\chi) = \underline{\omega}(\chi_1)^{\otimes a_1} \otimes \cdots \otimes \underline{\omega}(\chi_g)^{\otimes a_g}. \tag{1.16}$$

It is a line bundle over $\mathfrak{M}(B, \mu_N)$. We remark that if $\underline{\omega}$ is defined by a cocycle $\xi \in H^1(\mathfrak{M}(B, \mu_N), (\mathcal{O}_L \otimes \mathcal{O}_{\mathfrak{M}(B,\mu_N)})^\times)$ then $\underline{\omega}(\chi)$ is defined by the image of ξ under the map

$$H^1(\mathfrak{M}(B, \mu_N), (\mathcal{O}_L \otimes \mathcal{O}_{\mathfrak{M}(B,\mu_N)})^\times) \xrightarrow{\chi} H^1(\mathfrak{M}(B, \mu_N), \mathcal{O}^\times_{\mathfrak{M}(B,\mu_N)}) \ . \tag{1.17}$$

Note that the same construction applies for every abelian scheme $\underline{A} \longrightarrow S$ (where $S \longrightarrow \mathrm{Spec}(\mathcal{O}_{\tilde{L}}[d_L^{-1}])$ with real multiplication and μ_N level structure. Indeed, via the classifying morphism $S \longrightarrow \mathfrak{M}(\mu_N)$, $\underline{A}$ *is* $\underline{A}^S$. If we wish to make the dependence on $\underline{A}$ clear, we shall write $\underline{\omega}_{\underline{A}}(\chi)$.

Definition 1.1. A *Hilbert modular form* f over B, of μ_N-level and weight $\chi \in X(\mathbb{T}_B)$ is a section of $\underline{\omega}_B(\chi)$.

Equivalently,

Definition 1.2. A *Hilbert modular form* f over B, of μ_N-level and weight $\chi \in X(\mathbb{T}_B)$ is a rule

$$\underline{A} \mapsto f(\underline{A}) \in \underline{\omega}_{\underline{A}}(\chi), \tag{1.18}$$

commuting with base change and depending only on the isomorphism class of $\underline{A}$.

Let R be an $\mathcal{O}_{\tilde{L}}[d_L^{-1}]$-algebra and assume that $\underline{\omega}_{\underline{A}}$ is a free $\mathcal{O}_L \otimes R$ module of rank 1. A generator $\omega \in \underline{\omega}_{\underline{A}}$ is called a non-vanishing differential. We may equivalently define a Hilbert modular form as follows:

Definition 1.3. A *Hilbert modular form* f over B, of μ_N-level and weight $\chi \in X(\mathbb{T}_B)$ is a rule

$$(\underline{A}, \omega)_{/R} \mapsto f(\underline{A}, \omega) \in R \tag{1.19}$$

(R a B-algebra, ω a non-vanishing differential), commuting with base change and depending only on the isomorphism class of $(\underline{A}, \omega)_{/R}$, such that

$$f(\underline{A}, \alpha^{-1}\omega) = \chi(\alpha) f(\underline{A}, \omega), \quad \alpha \in (\mathcal{O}_L \otimes R)^\times = \mathbb{T}_B(R). \tag{1.20}$$

Remark 1.4. If $B = \mathbb{C}$, a Hilbert modular form of weight $\chi = \chi_1^{a_1} \cdots \chi_g^{a_g}$ is a complex Hilbert modular form of weight $(a_1, \ldots, a_g)$. See Chapter 3, Page 60. As a matter of notation we use

$$\mathrm{Norm} = \chi_1 \cdots \chi_g. \tag{1.21}$$

Thus, a Hilbert modular form of parallel weight k is of weight Norm^k is this terminology.

Definition 1.5. We denote by

$$\mathcal{M}(B, \chi, \mu_N) \tag{1.22}$$

the space of Hilbert modular forms over B of weight χ and level μ_N.

2. Tate Objects and the q-expansion

The purpose of this section is to define and state the properties of the q-expansion of a Hilbert modular form over arbitrary base rings. This q-expansion is to be obtained as a "special value" of the Hilbert modular form at particular abelian varieties with real multiplication - The *Tate objects*. For simplicity of exposition we shall treat only the case of parallel weight. The analogous assertions for non-parallel weight are correct with the necessary evident modifications. E.g., the base rings should be over the ring of integers of a Galois closure of L (so that the different weights are defined, etc.).

Given a closed point $x \in \mathfrak{M}(B, \mu_N)$ and a choice of local parameters $t_1, \ldots, t_g$ at x such that $\widehat{\mathcal{O}_x} \cong B[[t_1, \ldots, t_g]]$ and a trivialization of $\mathfrak{t}^*_{\underline{A}^U/\mathfrak{M}(B,\mu_N)}$ around x, every modular form can be expressed uniquely as an element of $\widehat{\mathcal{O}_x}$ by a Taylor series expansion. Moreover, this expansion is the *value* of the modular form at $\underline{A}^U|_{\widehat{\mathcal{O}_x}}$. The q-expansion is obtained by performing a similar process around a cusp. That is, in its essence it is nothing more than a Taylor expansion. Alas, several problems arise:

- The cusps are not regular points of $\mathfrak{M}^*(B, \mu_N)$.
- The universal family $\underline{A}^U$ over $\mathfrak{M}(B, \mu_N)$ does not extend to a family of *abelian varieties* over $\mathfrak{M}^*(B, \mu_N)$ and, in particular, there does not exists an abelian scheme with real multiplication over the completion of the local ring of a cusp that extrapolates the existing family over the "punctured" local ring.

Now, while for elliptic curves there is a totally satisfactory solution for this problem obtained by considering *generalized elliptic curves* (a family of generalized elliptic curves is allowed to have fibers that are a cyclic configuration of $\mathbb{P}^1$'s - an "N-gon") – a concept flexible enough to yield a universal object $\mathcal{E}^U/\mathfrak{M}^*_{\mathbb{Q}}(B, \mu_N)$ that extends the universal family of elliptic curves $\mathcal{E}^U/\mathfrak{M}_{\mathbb{Q}}(B, \mu_N)$, when $g > 1$ no such solution seems to be available at this time. Instead we must consider a whole class of polarized semiabelian schemes with real multiplication $\mathcal{A} \longrightarrow \mathrm{Spec}(S)$ over base schemes S, such that S is a normal local noetherian ring; the family is abelian outside the unique closed point s_0; s_0 maps to a particular cusp of $\mathfrak{M}^*_L(B, \mu_N)$. We then *evaluate* any modular form (of parallel weight) at $\mathcal{A}|_{Spec(S)-\{s_0\}}$ and obtain a value lying in a suitable localization of S. This value turns out to be dependent only on the cusp to which s_0 maps and "descends" to an element of the completion of the local ring of that cusp on $\mathfrak{M}^*_L(B, \mu_N)$.

Fix a totally real field L of degree g over $\mathbb{Q}$ and let $\mathcal{O}_L$ be the ring of integers and $\mathcal{D}^{-1}_{L/\mathbb{Q}}$ the inverse different. Let $\mathfrak{c}$ be a fixed fractional ideal of L. We think of $\mathfrak{c}$, or rather on its class in $Cl(L)^+$, as defining a module with a notion of positivity and therefore a component of $\mathfrak{M}(B, \mu_N)$.

Let $\mathfrak{a}, \mathfrak{b}$ be two fractional ideal of L such that

$$\mathfrak{c} = \mathfrak{a}\mathfrak{b}^{-1}. \tag{2.1}$$

Recall (Chapter 2, Section 2.2) that over $\mathbb{C}$, $\mathfrak{c}$-polarized abelian varieties with real multiplication come from the lattice $\mathfrak{b} \oplus (\mathcal{D}_L\mathfrak{a})^{-1}$ via various embeddings

$$\mathfrak{b} \oplus (\mathcal{D}_L\mathfrak{a})^{-1} \longrightarrow \mathfrak{b} \cdot \tau \oplus (\mathcal{D}_L\mathfrak{a})^{-1} \cdot 1, \quad \tau \in \mathcal{H}^g. \tag{2.2}$$

Indeed, we computed that the polarization module is $(\mathcal{D}_L \cdot \mathfrak{b} \cdot (\mathcal{D}_L \mathfrak{a})^{-1})^{-1} = \mathfrak{c}$. The moduli space over $\mathbb{C}$ of $\mathfrak{c}$-polarized abelian varieties with real multiplication is isomorphic to [2]

$$\mathrm{SL}(\mathfrak{b} \oplus (\mathcal{D}_L \mathfrak{a})^{-1}) \backslash \mathcal{H}^g. \tag{2.3}$$

The group $\mathrm{SL}(\mathfrak{b} \oplus (\mathcal{D}_L \mathfrak{a})^{-1})$ contains the translations

$$\tau \mapsto \tau + \alpha, \quad \alpha \in (\mathcal{D}_L \mathfrak{a}\mathfrak{b})^{-1}. \tag{2.4}$$

Therefore, every modular form over $\mathbb{C}$ has a q-expansion with respect to the totally positive elements of $\mathfrak{a}\mathfrak{b} = ((\mathcal{D}_L \mathfrak{a}\mathfrak{b})^{-1})^{\vee}$:

$$f_{(\mathfrak{a},\mathfrak{b})_{\mathrm{an}}} = a_0 + \sum_{\alpha \in (\mathfrak{a}\mathfrak{b})^+} a_\alpha q^\alpha, \quad q^\alpha = e^{2\pi i \cdot \mathrm{Tr}(\alpha \cdot \tau)}. \tag{2.5}$$

The ring $\mathbb{Z}[[q^\alpha : \alpha = 0 \text{ or } \alpha \in (\mathfrak{a}\mathfrak{b})^+]]$ (where $q^0 = 1$ and $q^\alpha q^\beta = q^{\alpha+\beta}$) is not noetherian – a somewhat awkward situation for algebraic geometry. We thus consider "approximations" of this ring. Let

$$\Delta = \{\ell_1, \cdots, \ell_g\} \tag{2.6}$$

be a set of $\mathbb{Q}$-linear functionals on the g-dimensional rational vector space L that are independent and have the property

$$\ell_i(L^+) \subseteq \mathbb{Q}^+. \tag{2.7}$$

We say that an element $m \in L$ is Δ-*positive* if

$$\ell_i(m) \geq 0, \quad \forall i. \tag{2.8}$$

Given a lattice $M \subset L$ we let

$$M_{\Delta-\mathrm{pos}} = \{m \in M : m \text{ is } \Delta - \text{positive}\}. \tag{2.9}$$

One observes that $M_{\Delta-\mathrm{pos}}$ is a finitely generated monoid, whilst $M^+ = \{m \in M : m \in L^+\}$ is not. We also note that

$$M^+ = \bigcap_{\Delta} M_{\Delta-\mathrm{pos}}. \tag{2.10}$$

EXERCISE 2.1. Prove the assertions made on $M_{\Delta-\mathrm{pos}}$ and M^+. Prove that a discrete monoid $M \subset L$ is finitely generated iff $M \otimes \mathbb{Q}^+$ is finitely generated.

EXAMPLE 2.2. Let $L = \mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$. Let $M = \mathcal{O}_L$. Let ℓ_1 and ℓ_2 be the following functionals

$$\ell_1(a + b\sqrt{2}) = a, \quad \ell_2(a + b\sqrt{2}) = a - b. \tag{2.11}$$

Let $\Delta = \{\ell_1, \ell_2\}$. Note that $(\mathcal{O}_L)_{\Delta-\mathrm{pos}}$ is generated by $1 + \sqrt{2}$ and $-\sqrt{2}$. See Figure ****

Figure ****

[2]For any two ideals $\mathfrak{d}$, $\mathfrak{e}$ such that $(\mathcal{D}_L \mathfrak{d}\mathfrak{e})^{-1} = \mathfrak{c}$ we may take $\mathrm{SL}(\mathfrak{d} \oplus \mathfrak{e}) \backslash \mathcal{H}^g$. The advantage in choosing different groups is that the stabilizer of $i\infty$ in $\mathrm{SL}(\mathfrak{d} \oplus \mathfrak{e}\mathcal{H}^g)$ contains $\mathfrak{d}^{-1}\mathfrak{e}$. Thus, by the choosing properly the group we mod by we get rid of the need to work with other cusps beside $i\infty$.

Note that ${\mathcal{O}_L}^+$ is not finitely generated. See Figure ****

Figure ****

We define

$$\mathbb{Z}[[M;\Delta]] = \left\{ \sum a_\alpha q^\alpha : a_\alpha \in \mathbb{Z}, \alpha \in M_{\Delta-\mathrm{pos}} \right\}. \tag{2.12}$$

It is a ring under the rules

$$q^0 = 1, \quad q^\alpha q^\beta = q^{\alpha+\beta}. \tag{2.13}$$

Note that if $x_1, \dots, x_n$ are generators of $M_{\Delta-\mathrm{pos}}$ then $\mathbb{Z}[[M;\Delta]]$ is a quotient of $\mathbb{Z}[[x_1, \dots, x_n]]$. Therefore, $\mathbb{Z}[[M;\Delta]]$ is a local noetherian ring.

EXERCISE 2.3. Prove that $\mathbb{Z}[[M;\Delta]]$ is complete with respect to the ideal I generated by $\{q^\alpha : \alpha \in M^+\}$. Prove that it is a normal domain.

We let

$$\mathbb{Z}((M;\Delta)) = \mathbb{Z}[[M;\Delta]][U^{-1}], \quad U = \{q^\alpha : \alpha \in M^+\}. \tag{2.14}$$

EXERCISE 2.4. Let α be an element of M such that $\ell_i(\alpha) > 0$ for all i. Prove that

$$\mathbb{Z}((M;\Delta)) = \mathbb{Z}[[M;\Delta]][(q^\alpha)^{-1}]. \tag{2.15}$$

Prove also that

$$\mathbb{Z}((M;\Delta)) = \{\sum_\alpha a_\alpha q^\alpha : a_\alpha \in \mathbb{Z}, \alpha \in M, \exists n \text{ s.t. } a_\alpha = 0 \text{ unless } \ell_i(\alpha) \geq n, \forall i\}. \tag{2.16}$$

EXERCISE 2.5. Prove that the natural map $\alpha \mapsto \epsilon\alpha$ defines an action of ${\mathcal{O}_L}^{\times +}$ on the ring $\mathbb{Z}((M;\Delta))$. Note that it doesn't act on $\mathbb{Z}[[M;\Delta]]$.[3]

Consider the torus $\mathbb{T} = \mathbb{G}_m \otimes_{\mathbb{Z}} \mathcal{D}_{L/\mathbb{Q}}^{-1}\mathfrak{a}^{-1}$ over the ring $S = \mathbb{Z}((\mathfrak{a}\mathfrak{b};\Delta))$. Recall that for any $R \in \mathbf{S}-\mathbf{alg}$ we have

$$\mathbb{T}(R) = R^\times \otimes_{\mathbb{Z}} \mathcal{D}_{L/\mathbb{Q}}^{-1}\mathfrak{a}^{-1} \tag{2.17}$$

and that characterizes $\mathbb{T}$. The character group of $\mathbb{T}$ is $\mathfrak{a} = (\mathcal{D}_{L/\mathbb{Q}}^{-1}\mathfrak{a}^{-1})^\vee$.

[3]When dealing with non-parallel weight one needs to modify the action by introducing the $r_k(\epsilon)$ factors.

We claim that there exists a canonical $\mathcal{O}_L$-linear homomorphism

$$\underline{q} : \mathfrak{b} \longrightarrow \mathbb{G}_m \otimes_{\mathbb{Z}} \mathcal{D}_{L/\mathbb{Q}}^{-1}\mathfrak{a}^{-1}, \tag{2.18}$$

where $\mathfrak{b}$ can be interpreted as a constant group scheme. To begin with, there is a given $\mathbb{Z}$-linear map

$$q : \mathfrak{a}\mathfrak{b} \longrightarrow \mathbb{G}_m(S), \tag{2.19}$$

given by

$$q(\alpha) = q^{\alpha}. \tag{2.20}$$

We can define a unique $\mathcal{O}_L$-linear map,

$$\underline{q} : \mathfrak{a}\mathfrak{b} \longrightarrow \mathbb{G}_m \otimes_{\mathbb{Z}} \mathcal{D}_{L/\mathbb{Q}}^{-1}, \tag{2.21}$$

having the property

$$(1 \otimes \mathrm{Tr}_{L/\mathbb{Q}})(\underline{q}(\alpha)) = q(\alpha) = q^{\alpha}, \tag{2.22}$$

where of course

$$1 \otimes \mathrm{Tr}_{L/\mathbb{Q}} : \mathbb{G}_m \otimes_{\mathbb{Z}} \mathcal{D}_{L/\mathbb{Q}}^{-1} \longrightarrow \mathbb{G}_m \otimes_{\mathbb{Z}} \mathbb{Z} = \mathbb{G}_m. \tag{2.23}$$

Tensoring Equation (2.21) by $\otimes_{\mathcal{O}_L}\mathfrak{a}^{-1}$, we get the map

$$\underline{q} : \mathfrak{b} \longrightarrow \mathbb{G}_m \otimes_{\mathbb{Z}} \mathcal{D}_{L/\mathbb{Q}}^{-1}\mathfrak{a}^{-1}. \tag{2.24}$$

Fact 2.6. The quotient $\mathbb{G}_m \otimes \mathcal{D}_{L/\mathbb{Q}}^{-1}\mathfrak{a}^{-1}/\underline{q}(\mathfrak{b})$ can be algebraized to an abelian scheme with real multiplication (called *Tate objects*) by $\mathcal{O}_L$, $\mathbf{Tate}_{\mathfrak{a},\mathfrak{b}}(\underline{q})$, over the ring $S = \mathbb{Z}((\mathfrak{a}\mathfrak{b}; \Delta))$ which carries a canonical $\mathfrak{c} = \mathfrak{a}\mathfrak{b}^{-1}$-polarization. The relative cotangent space is given by

$$\mathfrak{t}^*_{\mathbf{Tate}_{\mathfrak{a},\mathfrak{b}}(\underline{q})/S} \cong S\frac{dt}{t} \otimes_{\mathbb{Z}} \mathfrak{a}. \tag{2.25}$$

If $\mathfrak{a} = \mathcal{O}_L$ (and hence $\mathfrak{b} = \mathfrak{c}^{-1}$) the abelian scheme $\mathbf{Tate}_{\mathfrak{a},\mathfrak{b}}(\underline{q})$ carries a canonical differential

$$\omega_{can} = \frac{dt}{t} \otimes 1, \tag{2.26}$$

and a canonical μ_N-level for every N

$$\beta_{N,can} : \mu_N \otimes \mathcal{D}_{L/\mathbb{Q}}^{-1} \hookrightarrow \mathbb{G}_m \otimes \mathcal{D}_{L/\mathbb{Q}}^{-1}/\underline{q}(\mathfrak{c}^{-1}). \tag{2.27}$$

Definition 2.7. The Tate objects $\mathbf{Tate}_{\mathcal{O}_L,\mathfrak{c}^{-1}}(\underline{q})$ over $S = \mathbb{Z}((\mathfrak{c}^{-1}; \Delta))$ are called *standard Tate objects* (or *standard cusps*). They have a canonical $\mathfrak{c}$-polarization, non-vanishing differential ω_{can} and canonical μ_N-level structures $\beta_{N,can}$ for every N. We shall denote such an object by $\mathbf{Tate}_{\mathfrak{c}}(\underline{q})$.

We note that

$$\mathbf{Tate}_{\mathfrak{a},\mathfrak{b}}(\underline{q}) \otimes_{\mathcal{O}_L} \mathfrak{c} \cong \mathbb{G}_m \otimes_{\mathbb{Z}} \mathcal{D}_{L/\mathbb{Q}}^{-1}\mathfrak{a}^{-1}\mathfrak{a}\mathfrak{b}^{-1}/\underline{q}(\mathfrak{b}\mathfrak{a}\mathfrak{b}^{-1}) \tag{2.28}$$

$$\cong \mathbb{G}_m \otimes \mathcal{D}_{L/\mathbb{Q}}^{-1}\mathfrak{b}^{-1}/\underline{q}(\mathfrak{a}) \tag{2.29}$$

$$= \mathbf{Tate}_{\mathfrak{b},\mathfrak{a}}(\underline{q}). \tag{2.30}$$

That is, $\mathbf{Tate}_{\mathfrak{b},\mathfrak{a}}(\underline{q})$ is the dual abelian variety to $\mathbf{Tate}_{\mathfrak{a},\mathfrak{b}}(\underline{q})$.

Given a Hilbert modular form (of parallel weight) over a base ring B, we define its *q-expansion* at the standard cusp $\mathbf{Tate}_{\mathfrak{a},\mathfrak{b}}(\underline{q})$ by

(2.31)
$$f_{\mathfrak{c}}(q) = f(\mathbf{Tate}_{\mathfrak{c}}(\underline{q}) \otimes_S S_B, \iota_{can}, \lambda_{can}, \omega_{can}, \beta_{N,can}) \in S_B := \mathbb{Z}((\mathfrak{c}^{-1}; \Delta)) \otimes_{\mathbb{Z}} B.$$

In general, to obtain q-expansions at every cusp we need to *specify* a differential and level structures. Thus given B, one chooses an isomorphism

$$j : \mathfrak{a} \otimes_{\mathbb{Z}} B \cong \mathcal{O}_L \otimes_Z B. \tag{2.32}$$

This allows one to identify $\mathbb{G}_m \otimes \mathcal{D}_{L/\mathbb{Q}}^{-1}\mathfrak{a}^{-1}$ with $\mathbb{G}_m \otimes \mathcal{D}_{L/\mathbb{Q}}^{-1}$ over the base $S_B = \mathbb{Z}((\mathfrak{c}^{-1}; \Delta)) \otimes_{\mathbb{Z}} B$. We let $\omega(j)$ and $\beta_{N,j}$ be the resulting non-vanishing differential and μ_N-level structures obtained via this identification.

Definition 2.8. The q-expansion of a Hilbert modular form f defined over B (of parallel weight) at the cusp $\mathbf{Tate}_{\mathfrak{a},\mathfrak{b},j}(\underline{q})$ is

(2.33)
$$f_{\mathfrak{a},\mathfrak{b},j}(\underline{q}) = f(\mathbf{Tate}_{\mathfrak{a},\mathfrak{b},j}(\underline{q})_{/S_B}, \iota_{can}, \lambda_{can}, \omega(j), \beta_{N,j}) \in S_B := \mathbb{Z}((\mathfrak{a}\mathfrak{b}; \Delta)) \otimes_{\mathbb{Z}} B.$$

Note that immediately from the definition we obtain that the q-expansion commutes with base change and has bounded denominators.

Theorem 2.9. (q-expansion principle) *Let $g > 1$.* [4]

1. *The q-expansion $f_{\mathfrak{a},\mathfrak{b},j}(q)$ of a Hilbert modular form f defined over B, of polarization module $\mathfrak{c} = \mathfrak{a}\mathfrak{b}^{-1}$ and of level N at the cusp $\mathbf{Tate}_{\mathfrak{a},\mathfrak{b},j}(\underline{q})$ is independent of Δ and lies in fact in $\mathbb{Z}((\ (\mathfrak{a}\mathfrak{b})^+\ ))^{U_N^2} \otimes B$, where U_N is the group of units of $\mathcal{O}_L$ that are congruent to 1 modulo N. The ring $\mathbb{Z}((\ (\mathfrak{a}\mathfrak{b})^+\ ))^{U_N^2} \otimes B$ is isomorphic to the completion of the local ring of the cusp $(\mathfrak{a}, \mathfrak{b}, j)$ on $\mathfrak{M}^*(B, \mu_N)$.*
2. *The q-expansion map is an injective Galois equivariant homomorphism*

$$\mathcal{M}(B, \mathrm{Norm}^k, \mu_N) \longrightarrow \mathbb{Z}((\ (\mathfrak{a}\mathfrak{b})^+\ ))^{U_N^2} \otimes B \tag{2.34}$$

   that commutes with base change.
3. *Let f be a complex Hilbert modular form of polarization module $\mathfrak{c}$. Let $j_{\mathbb{C}} : \mathfrak{a} \otimes \mathbb{C} \longrightarrow \mathcal{O}_L \otimes \mathbb{C}$ be the natural identification. Then the algebraic q-expansion*

$$f_{\mathfrak{a},\mathfrak{b},j}(q) = f(\mathbf{Tate}_{\mathfrak{a},\mathfrak{b},j}(\underline{q})_{/S_B}, \iota_{can}, \lambda_{can}, \omega(j), \beta_{N,j}) \tag{2.35}$$

   is equal to the analytic q-expansion

$$f_{(\mathfrak{a},\mathfrak{b})_{an}} = a_0 + \sum_{\alpha \in (\mathfrak{a}\mathfrak{b})^+} a_\alpha q^\alpha. \tag{2.36}$$

We refer the reader to [**9**], [**97**], [**60**], [**31**] and [**30**] for a complete exposition of the theory of q-expansions.

[4] For $g = 1$ see Chapter 4, Theorem 3.5. The only difference is the Koecher principle.

3. Hasse Invariants

3.1. Definition and main properties of partial Hasse invariants. Let R be a ring of characteristic p. Let $f : A \longrightarrow \mathrm{Spec}(R)$ be an abelian scheme. The exact sequence for R a field

$$0 \longrightarrow H^0(A, \Omega^1_{A/R}) \longrightarrow H^1_{dR}(A/R) \longrightarrow H^1(A, \mathcal{O}_A) \longrightarrow 0 \tag{3.1}$$

can be jazzed up to a sequence of sheaves [5]

$$0 \longrightarrow R^0 f_* \Omega^1_{A/R} \longrightarrow H^1_{dR}(A/R) \longrightarrow R^1 f_* \mathcal{O}_A \longrightarrow 0. \tag{3.5}$$

The identification $H^1(A, \mathcal{O}_A) \cong \mathfrak{t}_{A^t/R}$ for R a field, where $\mathfrak{t}_{A^t/R}$ is the tangent space at zero $T_{A^t,0}$ or equivalently $\mathrm{Lie}(A^t/R)$ [6], can be extended to an identification

$$R^1 f_* \mathcal{O}_A \cong \mathfrak{t}_{A^t/R}. \tag{3.6}$$

Given a polarization $\lambda : A \longrightarrow A^t$ of degree prime to p we get an isomorphism

$$\lambda_* : \mathfrak{t}_{A/R} \longrightarrow \mathfrak{t}_{A^t/R}, \tag{3.7}$$

and hence, a perfect paring

$$< \cdot, \cdot >_\lambda \colon R^0 f_* \Omega^1_A \times R^1 f_* \mathcal{O}_A \longrightarrow R, \quad (a, b) \mapsto < a, \lambda_*^{-1}(b) > . \tag{3.8}$$

Here $< a, \lambda_*^{-1}(b) >$ stands for the natural pairing

$$\mathfrak{t}^*_{A/R} \times \mathfrak{t}_{A/R} \longrightarrow R. \tag{3.9}$$

Assume that p is not ramified in L. Let $\mathbb{F}$ be a fixed finite field isomorphic to the residue field of any of the residue fields of the prime factors of p in $\tilde{L}$. We now define modular forms $h_1, \ldots, h_g$ in characteristic p (more precisely, over $\mathbb{F}$) that we call *partial Hasse invariants*.

Assume that $\underline{A} = (A, \iota, \beta_N)_{/R}$ is an abelian variety with real multiplication and μ_N-level over an $\mathbb{F}$-algebra R. Assume further that $\mathfrak{t}^*_{A/R}$ is a free $\mathcal{O}_L \otimes R$ module. Let ω be a non-vanishing differential. Then ω gives an R-basis to $\mathfrak{t}^*_{A/R}$

$$\omega_1, \ldots, \omega_g. \tag{3.10}$$

Letting e_i be the i-th idempotent of the ring $\mathcal{O}_L \otimes R = \oplus_{i=1}^g R$, we have $\omega_i = e_i \omega$. Let λ be an $\mathcal{O}_L$-linear polarization of degree prime to p.

LEMMA 3.1. *Such λ exists.*

[5] Here $H^i_{dR}(A/R)$ are the right derived hypercohomology $R^i f_* \Omega^*_{A/R}$. That is, we have the hyper cohomology of the global sections functor with respect to the sequence

$$0 \longrightarrow *\mathcal{O}_A \xrightarrow{d} *\Omega^1_{A/R} \xrightarrow{d} *\Omega^2_{A/R} \xrightarrow{d} \cdots \tag{3.2}$$

($\Omega^i_{A/R} = (\Omega_{A/R})^{\otimes i}$). See [**44**, Section 11.4], and [**87**, Section 5]. It comes with a spectral sequence

$$H^q(A, \Omega^p_{A/R}) \Rightarrow H^n_{dR}(A/R). \tag{3.3}$$

The construction can be made into a sheaf on the base. Furthermore, when the base is affine, as we indeed assume, the global sections of the sequence (3.5) give an exact sequence:

$$0 \longrightarrow H^0(A, \Omega^1_{A/R}) \longrightarrow H^1_{dr}(A/R) \longrightarrow H^1(A, \mathcal{O}_A) \longrightarrow 0. \tag{3.4}$$

[6] Note that over $\mathbb{C}$, the sequence $0 \longrightarrow \mathbb{Z} \longrightarrow \mathcal{O}_A \longrightarrow \mathcal{O}_A^\times \longrightarrow 0$ gives $H^1(A, \mathbb{Z}) \longrightarrow H^1(A, \mathcal{O}_A) \longrightarrow H^1(A, \mathcal{O}_A^\times)^0 = A^t$, making the identification $H^1(A, \mathcal{O}_A)$ clear in this case.

PROOF. This follows from the proof of Lemma 5.2 in Chapter 3. □

Let

$$< \cdot, \cdot >_\lambda : \mathfrak{t}^*_{A/R} \times \mathfrak{t}_{A^t/R} \longrightarrow R \tag{3.11}$$

be the associated pairing, and let

$$\eta_1, \ldots, \eta_g \tag{3.12}$$

be the dual basis over R to $\mathfrak{t}_{A^t/R} = R^1 f_* \mathcal{O}_A$.

LEMMA 3.2. *The natural action of* $\mathrm{Fr} : R^1 f_* \mathcal{O}_A \longrightarrow R^1 f_* \mathcal{O}_A$, *induced form* $\mathcal{O}_A \longrightarrow \mathcal{O}_A$ *by* $x \mapsto x^p$, *permutes the* $\mathcal{O}_L$*-eigenspaces* $R\eta_i$.

PROOF. It is enough to prove that $[a] \cdot \mathrm{Fr}(\eta_{i-1}) = \sigma_i(a)\mathrm{Fr}(\eta_{i-1})$ for $a \in \mathcal{O}_L$. But the action of Fr commutes with endomorphisms [7]. Thus

$$[a] \cdot \mathrm{Fr}(\eta_{i-1}) = \mathrm{Fr}([a] \cdot \eta_{i-1}) \tag{3.13}$$

$$= \mathrm{Fr}(\sigma_{i-1}(a)\eta_{i-1}) \tag{3.14}$$

$$= \sigma^p_{i-1}(a)\mathrm{Fr}(\eta_{i-1}) \tag{3.15}$$

$$= \sigma_i(a)\mathrm{Fr}(\eta_{i-1}). \tag{3.16}$$

□

Let $i \mapsto \phi(i)$ be the permutation induced by the action of Frobenius on $R\eta_i$.

DEFINITION 3.3. The i-th partial Hasse invariant $h_i(\underline{A}, \omega)$ is the unique element of R such that

$$\mathrm{Fr}(\eta_{\phi^{-1}(i)}) = h_i(\underline{A}, \omega)\eta_i. \tag{3.17}$$

We shall write symbolically

$$h_i(\underline{A}, \omega) = \frac{\mathrm{Fr}(\eta_{\phi^{-1}(i)})}{\eta_i}. \tag{3.18}$$

THEOREM 3.4. *The rule*

$$(\underline{A}, \omega) \mapsto h_i(\underline{A}, \omega) \tag{3.19}$$

is a Hilbert modular form over $\mathbb{F}$, *of level* 1 *and weight* $\chi^p_{\phi^{-1}(i)}\chi_i^{-1}$ *and is independent of* λ. *Its* q*-expansion at every cusp is* 1.

PROOF. To simplify notation we shall assume henceforth that p is inert in $\mathcal{O}_{\tilde{L}}$. In this case we may work over $\mathcal{O}_{L,p}$-algebras. We choose $\mathbb{F} = \mathcal{O}_L/(p)$. We write $\sigma_1, \ldots, \sigma_g$ for the embedding $L \longrightarrow L_p$. We assume that

$$\sigma \circ \sigma_i = \sigma_{i+1}, \tag{3.20}$$

where σ is the unique lift of the Frobenius map $\mathrm{Fr} : \mathbb{F} \longrightarrow \mathbb{F}$ given by $x \mapsto x^p$.

The verification that the rule h_i commutes with base change and depends only on the isomorphism class is straightforward. It is also evidently independent of the μ_N-level structure β_N. We calculate the weight:

Let $\alpha \in (\mathcal{O}_L \otimes R)^\times$. The basis for $\mathfrak{t}^*_{A/R}$ obtained from $\alpha^{-1}\omega$ is

$$e_i([\alpha^{-1}]\omega) = [e_i \cdot \alpha^{-1}] \cdot [e_i] \cdot \omega = \chi_i(\alpha)^{-1}[e_i]\omega, \quad i = 1, \cdots, g. \tag{3.21}$$

[7] Either because of the fact that endomorphisms induce homomorphisms of Dieudonné modules, or because we are using two *independent* factorialities of $H^1(A, \mathcal{O}_A)$: with respect to the space A (endomorphisms) and with respect to the coefficients $\mathcal{O}_A$ (Frobenius).

The dual basis is thus

$$\chi_1(\alpha)\eta_1, \cdots, \chi_g(\alpha)\eta_g, \tag{3.22}$$

and

$$\frac{\mathrm{Fr}(\chi_{i-1}(\alpha)\eta_{i-1})}{\chi_i(\alpha)\eta_i} = \chi_{i-1}^p(\alpha)\chi_i^{-1}(\alpha) \cdot \frac{\mathrm{Fr}(\eta_{i-1})}{\eta_i}. \tag{3.23}$$

That is,

$$h_i(\underline{A}, \alpha^{-1}\omega) = (\chi_{i-1}^p\chi_i^{-1})(\alpha)h_i(\underline{A}, \omega). \tag{3.24}$$

Hence, h_i is of weight $\chi_{i-1}^p\chi_i^{-1}$.

Let λ' be another $\mathcal{O}_L$-linear polarization of degree prime to p. Then $\lambda' = m\lambda$ for some $m \in L$ such that $(m, p) = 1$. We have

$$< x, y >_\lambda = < x, my >_{\lambda'} . \tag{3.25}$$

Hence, if $\eta_1, \ldots, \eta_g$ are λ dual to $\omega_1, \cdots, \omega_g$ then $m\eta_1, \cdots, m\eta_g$ are λ' dual to $\omega_1, \cdots, \omega_g$. But,

$$\frac{\mathrm{Fr}(m\eta_{i-1})}{m\eta_{i-1}} = \frac{m\mathrm{Fr}(\eta_{i-1})}{m\eta_{i-1}} = \frac{\mathrm{Fr}(\eta_{i-1})}{\eta_{i-1}}. \tag{3.26}$$

Therefore, h_i is independent of λ.

We now address the issue of q-expansions. We first remark that it would follow from the map

$$r : \bigoplus_{\chi \in X(\mathbb{T})} \mathcal{M}(\mathbb{F}, \chi, \mu_N) \longrightarrow R, \tag{3.27}$$

(which is the analogue of the map (5.10) in Theorem 4.1) that it is enough to prove that the q-expansion of h_i is 1 at *one* cusp of *every* component of the moduli space $\mathfrak{M}(\mathbb{F}, \mu_N)$. Let $\mathbf{Tate} = \mathbb{G}_m \otimes \mathcal{D}_L^{-1}/q(\mathfrak{b})$ be a standard Tate object with its canonical $\mathcal{O}_L$-structure ι_{can}, canonical μ_N-level structure $\beta_{N,can}$, and canonical non-vanishing differential ω_{can}. It is a scheme over S where S *** Then

$$\mathfrak{t}^*_{\mathbf{Tate}/S} \cong S \otimes_{\mathbb{Z}} \mathcal{O}_L \cdot \frac{dt}{t}. \tag{3.28}$$

The canonical isomorphism is provided by the invariant differential dt/t of $\mathbb{G}_m$ with coordinate t.

We have, over a perfect ring of characteristic p:

$$< \mathrm{Fr}(\eta_{i-1}), \omega_i >_\lambda = < \eta_{i-1}, \mathrm{Ver}(\omega_i) >_\lambda^{(1/p)} . \tag{3.29}$$

Therefore, it is enough to show that

$$\mathrm{Ver}(w_{can,i}) = w_{can,i-1}, \quad \forall i. \tag{3.30}$$

This amounts to the identity

$$e_i\mathrm{Ver} = \mathrm{Ver}\, e_{i-1} \tag{3.31}$$

as operators on $\mathfrak{t}^*_{\mathbf{Tate}/S}$. Now,

$$(a\otimes)1\mathrm{Ver} = \mathrm{Ver}(a \otimes 1), \quad (a \otimes r^p)\mathrm{Ver} = \mathrm{Ver}(1 \otimes r). \tag{3.32}$$

Hence, we just assert that the automorphism $v : \mathcal{O}_L \otimes R \longrightarrow \mathcal{O}_L \otimes R$ given by $v(a \otimes r) = a \otimes r^p$ takes e_{i-1} to e_i. Since clearly $v(e_{i-1})$ is a multiple of e_i (consider the action of $\mathcal{O}_L$) the assertion follows from e_i being orthogonal idempotents

$$1 = e_1 + \cdots + e_g, \quad e_i^2 = e_i, \ \ \forall i. \tag{3.33}$$

□

For further study of the partial Hasse invariants see [**37**]. In Corollary 11.7 we prove the following

THEOREM 3.5. *The divisor of h_i is reduced.*

Let

$$H = h_1 \cdots h_g \tag{3.34}$$

be the total Hasse invariant. It is the determinant of the Hasse-Witt matrix and is a Hilbert modular form over $\mathbb{F}$ (in fact over $\mathbb{Z}/p\mathbb{Z}$) of level 1 and of weight Norm^{p-1}. It has q-expansion 1 at every cusp. In Chapter 11.2 we relate the divisors of the h_i's to a certain stratification of $\mathfrak{M}(\mathbb{F}, \mu_N)$ studied in [**39**]. In particular it follows that the divisors

$$W_i = (h_i), \tag{3.35}$$

are reduced and "as transversal to each other as possible".

EXERCISE 3.6. Prove that for $g = 1$ and $p \geq 5$, H is the reduction modulo p of the normalized Eisenstein series E_{p-1}. What goes wrong for $p = 2, 3$? Can you provide a geometric explanation for this?

3.2. Further properties. Following [**116**] we state prove the following

LEMMA 3.7. *Let $\underline{A} = (A, \iota, \beta_N)_{/k}$ be an ordinary abelian variety with real multiplication (and μ_N-level) over a field $k \supseteq \mathbb{F}$. Then $A[p]^0 \cong \mu_p \otimes \mathcal{D}_L^{-1}$ iff there exists a non-vanishing differential $\omega \in \Omega^1_{A/k} = \Omega^1_{A[p]^0/k}$ such that $h_i(\underline{A}, \omega) = 1$ for all i.*

PROOF. If $A[p]^0 \cong \mu_p \otimes \mathcal{D}_L^{-1}$ then the same computation giving the q-expansion shows that for the canonical non-vanishing differential ω on $\mu_p \otimes \mathcal{D}_L^{-1}$ we get $h_i(\underline{A}, \omega) = 1$ for every i.

Conversely, let $\omega \in \Omega^1_{A/k}$ be a non-vanishing differential such that $h_i(\underline{A}, \omega) = 1$ for every i. We may assume without loss of generality that k is a perfect field. We identify $H^0(A, \Omega^1_{A/k})$ with the Dieudonné module of $A[p]^0$. The Dieudonné module of $\mu_p \otimes \mathcal{D}_{L/\mathbb{Q}}^{-1}$ is canonically $k \otimes \mathcal{O}_L$, where the $\mathcal{O}_L$ action is the evident one: Fr is the zero map and Ver is the map determined by $x^p \otimes m \mapsto x \otimes m$.

We define a map

$$\psi : H^0(A, \Omega^1_{A/k}) \longrightarrow k \otimes \mathcal{O}_L \tag{3.36}$$

by

$$(\sum x_i \otimes m_i)\omega \mapsto \sum x_i \otimes m_i. \tag{3.37}$$

It is obviously an isomorphism of $k \otimes \mathcal{O}_L$ modules. Since Fr is zero on $H^0(A, \Omega^1_A)$ the only question remaining is whether ψ is Ver-equivariant. But this just boils down to the assertion that $\mathrm{Ver}(e_i\omega) = e_{i-1}\omega$, which in turn holds iff $h_i(\underline{A}, \omega) = 1$. □

4. The Kernel of the q-expansion

Similar to the case of $g = 1$, the identification of the kernel of the q-expansion map is based on comparing *modular forms* of μ_N-level with *functions* of μ_{Np} level. To simplify the exposition, *we assume that p is inert in L.* This is merely a technical requirement if one assumes that p is unramified in L. See for [**37**] an explanation of the yoga.

Let $N \geq 4$ be a fixed integer prime to p. As before, we denote by $\mathfrak{M}(\mathbb{F}, \mu_N)$ the fine moduli scheme of triples $\underline{A} = (A, \iota, \beta_N)_{/S}$ of abelian varieties A with real multiplication ι and μ_N level structure $\beta_N : \mu_N \otimes \mathcal{D}_L^{-1} \hookrightarrow A$ over $\mathbb{F}$ schemes S. We let $\mathfrak{M}^*(\mathbb{F}, \mu_N)$ be the compactification obtained by adjoining the cusps. We let $\mathfrak{M}(\mathbb{F}, \mu_{Np})$ (resp. $\mathfrak{M}^*(\mathbb{F}, \mu_{Np})$) be the moduli space with μ_{Np}-structure, which we write as $\beta_N \times \beta_p$ (resp. with the cusps added). The morphisms

$$\mathfrak{M}(\mathbb{F}, \mu_{Np}) \longrightarrow \mathfrak{M}(\mathbb{F}, \mu_N)^{\mathrm{ord}}, \quad \mathfrak{M}^*(\mathbb{F}, \mu_{Np}) \longrightarrow \mathfrak{M}^*(\mathbb{F}, \mu_N)^{\mathrm{ord}}, \tag{4.1}$$

are Galois coverings (with Galois group isomorphic to $(\mathcal{O}_L/(p))^\times$) of the open scheme consisting of the ordinary locus of $\mathfrak{M}(\mathbb{F}, \mu_N)$ and $\mathfrak{M}^*(\mathbb{F}, \mu_N)$, respectively (the cusps are ordinary). The components of $\mathfrak{M}^*(\mathbb{F}, \mu_{Np})$ are the same as those of $\mathfrak{M}^*(\mathbb{F}, \mu_N)^{\mathrm{ord}}$. This follows from Ribet's Theorem 6.19.

THEOREM 4.1. *Fix a polarization module. Let $(\mathfrak{a}, \mathfrak{a}^+)$ be a representative of the polarization module. Let* **Tate** *be a standard Tate object for this ideal $\mathfrak{a}$ as in Section 2. The kernel of the q-expansion map,*

$$\bigoplus_{\chi \in X(\mathbb{T}_\mathbb{F})} \mathcal{M}(\mathbb{F}, \chi, \mu_N) \longrightarrow \mathbb{F}[[q^\nu : \nu \in H]] \tag{4.2}$$

*(H consisting of the totally positive elements of ***), is the ideal*

$$(h_1 - 1, \ldots, h_g - 1). \tag{4.3}$$

PROOF. We consider **Tate** as a cusp on $\mathfrak{M}(\mathbb{F}, \mu_{Np})$. Let R be the ring of regular functions on $\mathfrak{M}(\mathbb{F}, \mu_{Np})$. We define a surjective map,

$$r : \bigoplus_{\chi \in X(\mathbb{T}_\mathbb{F})} \mathcal{M}(\mathbb{F}, \chi, \mu_N) \longrightarrow R, \tag{4.4}$$

such that the composition,

$$\bigoplus_{\chi \in X(\mathbb{T}_\mathbb{F})} \mathcal{M}(\mathbb{F}, \chi, \mu_N) \longrightarrow R \hookrightarrow \widehat{\mathcal{O}_{\mathfrak{M}(\mathbb{F}, \mu_{Np}), \mathbf{Tate}}}, \tag{4.5}$$

is the q-expansion map. We then prove that

$$\mathrm{Ker}(r) = (h_1 - 1, \ldots, h_g - 1). \tag{4.6}$$

This, together with the surjectiveness of r, would prove the Theorem and also give Corollary 4.4.

To define the map r we first need a construction.

LEMMA 4.2. *For every $\chi \in X(\mathbb{T}_\mathbb{F})$ there exists a canonical modular form $a(\chi)$ on $\mathfrak{M}^*(\mathbb{F}, \mu_{Np})$ such that:*

- *$a(\chi)$ has weight χ.*
- *$a(\chi)$ has q-expansion 1 at* **Tate**.
- *$a(\chi)a(\chi') = a(\chi\chi')$.*
- *$a(\chi)$ doesn't vanish.*

PROOF. Given $(\underline{A}, \beta_p)_{/R}$ over a ring R. we have an induced isomorphism

$$\beta_{p*}; \mathfrak{t}^*_{\mu_p \otimes \mathcal{D}_L^{-1}} \longrightarrow \mathfrak{t}^*_{A/R}. \tag{4.7}$$

The image of dt/t is a non-vanishing differential ω on A/R. We let

$$a(\chi_i) = e_i \omega. \tag{4.8}$$

For $\chi = \chi_1^{a_1} \cdots \chi_g^{a_g}$ we let

$$a(\chi) = a(\chi_1)^{a_1} \cdots a(\chi_g)^{a_g}. \tag{4.9}$$

The properties of the $a(\chi)$'s are easily verified. □

We now define

$$r(f) = \frac{f}{a(\chi)}, \quad f \in \mathcal{M}(\mathbb{F}, \chi, \mu_N), \tag{4.10}$$

and extend it by linearity to

$$r : \underset{\chi \in X(\mathbb{T}_{\mathbb{F}})}{\oplus} \mathcal{M}(\mathbb{F}, \chi, \mu_N) \longrightarrow R. \tag{4.11}$$

The properties of the $a(\chi)$'s ensure it is a ring homomorphism. Furthermore, the composition in (4.5) is the q-expansion. First note that the q-expansion of $r(f)$ is the same as that of f. Second, it follows from *** that the q-expansion of $r(f)$ with respect to a Tate object is non other than its expression in the local ring of the cusp corresponding to that Tate object.

At this point we may conclude:

- $I := \mathrm{Ker}(r)$ is the kernel of the q-expansion map.
- $I \supseteq (h_1 - 1, \ldots, h_g - 1)$.
- The q-expansion map is injective on every fixed weight, because r restricted to any $\mathcal{M}(\mathbb{F}, \chi, \mu_N)$ is injective!

The next ingredient is defining a
Galois action: Let $\mathbb{F}^\times = (\mathcal{O}_L/(p))^\times$ act on

- $f \in \mathcal{M}(\mathbb{F}, \chi, \mu_N)$ by $[\alpha]f = \chi(\alpha)f$.
- $g \in R$ by $([\alpha]g)(\underline{A}, \beta_p) = g(\underline{A}, \beta_p \circ (1 \otimes \alpha))$.

One verifies that r is equivariant for this Galois action and that in fact $\mathbb{F}^\times = Gal(\mathfrak{M}^*(\mathbb{F}, \mu_{Np})/\mathfrak{M}^*(\mathbb{F}, \mu_N)^{\mathrm{ord}})$. Let us note that $\mathbb{F}^\times \hookrightarrow (\mathbb{F} \otimes \mathbb{F})^\times = \mathbb{T}_{\mathbb{F}}(\mathbb{F})$. Therefore, every character $\chi \in X(\mathbb{T}_{\mathbb{F}})$ gives a homomorphism $\chi : \mathbb{F}^\times \longrightarrow \mathbb{F}^\times$. We let

$$X(1)_p = \{\chi \in X(\mathbb{T}_{\mathbb{F}}) : \chi(\mathbb{F}^\times) = 1\} \tag{4.12}$$

$$=< \chi_1^p \chi_2^{-1}, \ldots, \chi_i^p \chi_{i+1}^{-1}, \ldots, \chi_g^p \chi_1^{-1} > . \tag{4.13}$$

Kummer theory gives

$$R = \underset{\chi \in X/X(1)_p}{\oplus} R^\chi. \tag{4.14}$$

LEMMA 4.3. *We have the identity* $[\alpha]a(\chi) = \chi^{-1}(\alpha)a(\chi)$.

PROOF. (Of Lemma). It is enough to prove that

$$[\alpha]a(\chi_i) = \chi_i^{-1}(\alpha)a(\chi_i). \tag{4.15}$$

Now, $a(\chi_i)(\underline{A}, \beta_p) = e_i \beta_{p*} \frac{dt}{t} \otimes 1$. Note that $1 \otimes \alpha$ acts by $1 \otimes \alpha$ on the tangent space of $\mu_p \otimes \mathcal{D}_{L/\mathbb{Q}}^{-1}$ and $(1 \otimes \alpha)_*$ acts of the cotangent space by $(1 \otimes \alpha)^{-1}$ whenever this is defined. Thus,

$$[\alpha] a(\chi_i)(\underline{A}, \beta_p) = e_i (\beta_p \circ (1 \otimes \alpha))_* \frac{dt}{t} \otimes 1 \tag{4.16}$$

$$= \beta_{p*} e_i (1 \otimes \alpha)_* \frac{dt}{t} \otimes 1 \tag{4.17}$$

$$= \chi_i(\alpha)^{-1} \beta_{p*} e_i \frac{dt}{t} \otimes 1 \tag{4.18}$$

$$= \chi_i(\alpha)^{-1} a(\chi_i)(\underline{A}, \beta_p). \tag{4.19}$$

□

r is surjective: Given $g \in R^\chi$ let

$$f = g \cdot a(\chi) \cdot H^n, \quad n \gg 0. \tag{4.20}$$

Then f is a holomorphic modular form on $\mathfrak{M}^*(\mathbb{F}, \mu_{Np})$ which is $\mathbb{F}^\times$-invariant. Hence f is a meromorphic modular form on $\mathfrak{M}^*(\mathbb{F}, \mu_N)$, which is holomorphic if $n \gg 0$. Clearly $r(f) = g$.

$\mathrm{Ker}(r) = (h_1 - 1, \ldots, h_g - 1)$: Let $f_{\psi_1} + \cdots + f_{\psi_r}$ be in the kernel of r (with f_{ψ_i} of weight $\psi_i \in X(\mathbb{T}_\mathbb{F})$). We may assume without loss of generality that for $i \neq j$, $\psi_i \neq \psi_j$.

Replacing f_{ψ_i} by $f_{\psi_i} + f_{\psi_i}(h_j - 1)$ sufficiently many times and for appropriate i's and j's we may further assume that

$$i \neq j \Rightarrow \psi_i \neq \psi_j \pmod{X(1)_p}. \tag{4.21}$$

But since r is equivariant, every $r(f_{\psi_i})$ "falls" in a different eigenspace in the decomposition $R = \oplus_{\chi \in X/X(1)_p} R^\chi$. Thus every $r(f_{\psi_i}) = 0$. But r is injective on every $\mathcal{M}(\mathbb{F}, \chi, \mu_N)$. Thus, $f_{\psi_i} = 0$ for every i.

□

Corollary 4.4. *We have*

$$\bigoplus_{\chi \in X(\mathbb{T}_\mathbb{F})} \mathcal{M}(\mathbb{F}, \chi, \mu_N)/(h_1 - 1, \ldots, h_g - 1) \cong R. \tag{4.22}$$

Corollary 4.5. *The kernel of the q-expansion*

$$\bigoplus_{k \in \mathbb{Z}} \mathcal{M}(\mathbb{F}, \mathrm{Norm}^k, \mu_N) \longrightarrow \mathbb{F}[[q^\nu : \nu \in ***]] \tag{4.23}$$

is $(H - 1)$.

Remark 4.6. Theorem 4.1 and its Corollaries hold for any cusp.

5. Applications

Corollary 5.1. *The q-expansion map over* $\mathbb{C}$ *is injective.*

Proof. The kernel of the q expansion is $\mathrm{Gal}(L^{alg}/L)$ equivariant and hence the kernel is generated by modular forms with coefficients in L. Let $\sum_{i=1}^t f_{\psi_i} \in$

$\oplus\mathcal{M}(\mathbb{C},\chi,\mu_N)$ be such a sum of modular forms with q-expansion equal to zero. We assume that for $i\neq j$ we have $\psi_i\neq\psi_j$. Now, choose a (inert[8]) prime p such that

- The q-expansion of every f_{ψ_i} is p-integral and is non-zero modulo p.
- The weights $\psi_1,\ldots,\psi_t$ are non congruent modulo $X(1)_p$.

Then the reduction of the forms f_{ψ_i} modulo p is well defined, $f_{\psi_i}\not\equiv 0 \pmod p$ and the weights are distinct in $X/X(1)_p$. It follows from Theorem 4.1 that each f_{ψ_i} (mod p) is zero. A contradiction. □

Recall Siegel's construction. For $g>1$ and $k\geq 2$ (or $g=1$ and $k\geq 4$) there exists an Eisenstein series $E_k^*=E_k^{L,*}$ (and a normalized Eisenstein series $E_k=E_k^L$) of level 1, weight Norm^k and q-expansion

$$E_k^*=2^{-g}\zeta_L(1-k)+\sum_{\nu\in\mathcal{O}_L{}^+}c_{k,\nu}q^\nu,\quad E_k=2^g\zeta_L(1-k)^{-1}E_k^*, \tag{5.1}$$

where

$$c_{k,\nu}=\sum_{\mathcal{O}_L\supseteq\mathfrak{c}\supseteq(\nu)}\mathrm{Norm}(\mathfrak{c})^{k-1}. \tag{5.2}$$

Corollary 5.2. *We have the following bound on denominators:*

$$k\not\equiv 0 \pmod{p-1}\Rightarrow\zeta_L(1-k)\ \textit{is } p\ \textit{integral.} \tag{5.3}$$

Proof. Assume that p divides the denominator of $\zeta_L(1-k)$. Then we have the following congruence of q-expansions:

$$E_k(q)-1\equiv 0 \pmod p. \tag{5.4}$$

However, the ring $\oplus\mathcal{M}(\mathbb{F},\chi,\mu_N)/(h_1-1,\ldots,h_g-1)$ is $X/X(1)_p$ graded. Thus, the homogenous parts of the relation $E_k(q)-1\equiv 0 \pmod p$ also belong to $(h_1-1,\ldots,h_g-1)$. Since $1\notin(h_1-1,\ldots,h_g-1)$, it follows that $E_k(q)-1$ is a homogenous element. That is

$$\mathrm{Norm}^k\equiv\mathbf{1} \pmod{X(1)_p}. \tag{5.5}$$

That is, the map $\mathrm{Norm}^k:\mathbb{F}^\times\longrightarrow\mathbb{F}^\times$ is the trivial character. Hence $(p-1)|k$. □

Corollary 5.3. *We have the following congruences between values of zeta functions. Suppose that $k\not\equiv 0 \pmod{p-1}$ then*

$$k\equiv k' \pmod{p-1}\Rightarrow\zeta_L(1-k)\equiv\zeta_L(1-k') \pmod p. \tag{5.6}$$

Proof. Let $\alpha=2^{-g}(\zeta_L(1-k)-\zeta_L(1-k'))$. Consider the congruence

$$E_k^*-E_{k'}^*-\alpha=0 \pmod p. \tag{5.7}$$

Using a grading argument, we infer from $\mathrm{Norm}^k\equiv\mathrm{Norm}^{k'}\not\equiv\mathbf{1} \pmod{X(1)_p}$ that $\alpha\equiv 0 \pmod p$. That is, $\zeta_L(1-k)\equiv\zeta_L(1-k') \pmod p$. □

This method of obtaining bounds on the denominator of the zeta function of L and congruences between its values at negative integers can be generalized considerably to cover the cases $k\equiv 0 \pmod{p-1}$ and to give congruences modulo p^m. The idea, following Serre and Katz, is to consider the moduli space modulo p^m, and use the easy congruences on the higher coefficients of suitable Eisenstein series to obtain

[8]If a number field has one inert prime, it has infinitely many inert primes. We insist on inert primes only because we formulated the q-expansion kernel for inert primes. But, in fact, the argument easily generalizes to any totally real number field and non-ramified primes.

information on the constant term. The details may be found in [**38**]. We include here the results for completeness and for the reader's convenience. It seems they are sharp.

PROPOSITION 5.4. *Let p be inert in L. Let $k \geq 2$.*

1. Let $p \neq 2$, then if $k \equiv 0 \pmod{p-1}$

$$\mathrm{val}_p(\zeta_L(1-k)) \geq -1 - \mathrm{val}_p(k), \tag{5.8}$$

and $\zeta_L(1-k)$ is p-integral if $k \not\equiv 0 \pmod{p-1}$.

2. If $p = 2$, then

$$\mathrm{val}_2(\zeta_L(1-k)) \geq g - 2 - \mathrm{val}_2(k). \tag{5.9}$$

PROPOSITION 5.5. *Let p be inert in L. Let $k, k' \geq 2$ and $k \equiv k' \pmod{(p-1)p^m}$.*

1. If $k \not\equiv 0 \pmod{p-1}$ then

$$(1 - p^{g(k-1)})\zeta_L(1-k) \equiv (1 - p^{g(k'-1)})\zeta_L(1-k') \pmod{p^{m+1}}. \tag{5.10}$$

2. If $k \equiv 0 \pmod{p-1}$ but $p \neq 2$, then

$$(1 - p^{g(k-1)})\zeta_L(1-k) \equiv (1 - p^{g(k'-1)})\zeta_L(1-k') \pmod{p^{m-1-\mathrm{val}_p(k\cdot k')}}. \tag{5.11}$$

3. If $p = 2$, then

$$(1 - 2^{g(k-1)})\zeta_L(1-k) \equiv (1 - 2^{g(k'-1)})\zeta_L(1-k') \pmod{2^{m+g-2-\mathrm{val}_2(k\cdot k')}}. \tag{5.12}$$

PROPOSITION 5.6. *There exists a notion of* filtration *on Hilbert modular forms. Given a q-expansion $b(q)$ such that it is a q expansion of some Hilbert modular form, there exists a unique modular form f_0 such that the set of all modular forms with q-expansion $b(q)$ is the set*

$$\{f_0 \cdot \prod_{i=1}^{g} h_i^{a_i} : a_i \geq 0\}. \tag{5.13}$$

We call the weight of f_0 the filtration of the q-expansion $b(q)$.

PROOF. If f and g have the same q-expansion then $r(f) = r(g)$, and vice versa. We are given that $b(q)$ is a q-expansion of some Hilbert modular form of weight, say, χ. Let f' be a function on $\mathfrak{M}(\mathbb{F}, \mu_{Np})$ such that $f' \in R^\chi$ and in the local ring of the appropriate cusp $f' = b(q)$. Then all the meromorphic modular forms having q expansion $b(q)$ are of the form $f' \cdot a(\chi) \cdot \prod h_i^{a_i}$ where the $a_i \in \mathbb{Z}$. But the divisor of h_i is a reduced effective divisor W_i. Therefore, there is a choice $a_1^*, \ldots, a_g^*$ such that $f_0 = f' \cdot a(\chi) \cdot \prod h_i^{a_i^*}$ is holomorphic and non-vanishing on $(H) = W_1 \cup \cdots \cup W_g$. It follows that every other holomorphic form with the same q-expansion is a multiple $f_0 \cdot \prod_{i=1}^g h_i^{a_i}$ with $a_i \geq 0$. □

6. p-adic Hilbert Modular Forms

This section is modeled after Chapter 4, Section 6. Many of the definitions and proofs extend verbatim, with additional supporting arguments needed every once in a while. We shall therefore be brief and let the reader refer back to Chapter 6 whenever needed. For simplicity we shall discuss only Hilbert modular forms of parallel level. The totally real field is of course fixed and is denoted by L; the prime p is unramified and for simplicity of exposition assumed to be inert in L.

6.1. Test objects and overconvergent forms. We fix a Hilbert modular form with respect to L, denoted for heuristic reasons by E_ℓ with the following properties:

- E_ℓ has level 1.
- E_ℓ has q-expansion $E_\ell(q)$ with coefficients in $\mathcal{O}_{L_p}$;
- $E_\ell(q)$ is congruent to 1 modulo p.
- E_ℓ has weight Norm^ℓ and $(p-1)|\ell$.

REMARK 6.1. 1. Such a modular form always exists. Indeed, for $n \gg 0$ the modular form H^n over $\mathbb{F}$ lifts to a modular form of level 1 over $\mathcal{O}_{L_p}$.
2. Given Leopoldt's conjecture one can choose E_ℓ to be the Eisenstein series E_ℓ^* in (5.1).

We fix:

- B – a p-adic ring.
- r – an element of B.

DEFINITION 6.2. Let C be a p-adic ring which is a B-algebra. A *test object* of:

- level μ_N,
- growth condition r,
- over B,

is a quadruple:

$$(\underline{A}, \omega, \beta_N, Y)_{/C}, \tag{6.1}$$

such that:

- $\underline{A}/C$ is an abelian variety with real multiplication (A, ι) defined over C such that condition **(R)** holds;
- $\omega \in \mathfrak{t}^*_{\underline{A}/C}$ is a relative non-vanishing differential;
- $\beta_N : \mu_{N/C} \otimes \mathcal{D}_L^{-1} \hookrightarrow \underline{A}$ is an embedding of group schemes over $\mathrm{Spec}(C)$;
- $Y \in C$ satisfies $Y \cdot E_\ell(\underline{A}, \omega) = r$.

If $(\underline{A}, \omega, \beta_N, Y)_{/C}$ is a test object, so is $(\underline{A}, \lambda\omega, \beta_N, \mathrm{Norm}\lambda^\ell Y)_{/C}$ for $\lambda \in C^\times$. Indeed,

$$(\mathrm{Norm}\lambda^\ell Y) \cdot (E_\ell(\underline{A}, \lambda\omega)) = (\mathrm{Norm}\lambda^\ell Y)(\mathrm{Norm}\lambda^{-\ell}) E_\ell(\underline{A}, \omega) = r. \tag{6.2}$$

Equivalently, a test object could be thought of as

$$(\underline{A}, \beta_N, Y)_{/C}, \;\; Y \in (\det \mathfrak{t}^*_{\underline{A}/C})^{-\ell} \tag{6.3}$$

such that $Y \cdot E_\ell = r$, where E_ℓ is interpreted as a rule associating to $\underline{A}/C$ a section of $(\det \mathfrak{t}^*_{\underline{A}/C})^\ell$.

DEFINITION 6.3. A *p-adic Hilbert modular form* (à la Katz) of:

- weight Norm^k, $k \in \mathbb{Z}$,
- level μ_N,
- growth r,
- defined over B,

is a rule associating to $(\underline{A}, \omega, \beta_N, Y)_{/C}$ an element $f(\underline{A}, \omega, \beta_N, Y) \in C$ such that:

- $f(\underline{A}, \omega, \beta_N, Y)$ depends only on the isomorphism class of the test object $(\underline{A}, \omega, \beta_N, Y)_{/C}$;
- the rule f commutes with base change;

- for every $\lambda \in C^\times$ and every test object $(\underline{A}, \omega, \beta_N, Y)_{/C}$

$$f(\underline{A}, \lambda\omega, \beta_N, \mathrm{Norm}\lambda^\ell Y) = \mathrm{Norm}\lambda^{-k} f(\underline{A}, \omega, \beta_N, Y). \tag{6.4}$$

DEFINITION 6.4. The space of Hilbert modular forms over B, of μ_N-level, weight Norm^k and growth condition r is denoted by $\mathbb{F}(B, k, \mu_N; r)$. If $r \notin B^\times$, they are called *overconvergent* modular forms.

Again, the study of modular forms of growth $r \in B^\times$ amounts to the case $r = 1$. In that case we discard all abelian varieties having non-ordinary reduction.

Every classical modular form f of weight Norm^k, level N, over B defines a p-adic modular form (still denoted f) in $\mathbb{F}(B, k, \mu_N; r)$:

$$f(\underline{A}, \omega, \beta_N, Y)_{/C} := f(\underline{A}, \omega, \beta_N)_{/C}. \tag{6.5}$$

As an example of a truly p-adic modular form, consider

$$f(\underline{A}, \omega, \beta_N, Y) = Y. \tag{6.6}$$

By definition,

$$f(\underline{A}, \lambda\omega, \beta_N, \mathrm{Norm}\lambda^\ell Y) = \mathrm{Norm}\lambda^\ell Y = \mathrm{Norm}\lambda^\ell f(\underline{A}, \omega, \beta_N, Y). \tag{6.7}$$

Therefore, it is a modular form in $\mathbb{F}(B, \mathrm{Norm}^{-\ell}, \mu_N; r)$.

6.2. q-expansion for p-adic modular forms. If $E_\ell(\underline{A}, \omega)$ is invertible, then in any test object $(\underline{A}, \omega, \beta_N, Y)$ we must have $Y = r \cdot E_\ell(\underline{A}, \omega)^{-1}$. This applies in particular to the standard Tate object $\mathbf{Tate}_{\mathfrak{c}}(\underline{q}) = \mathbb{G}_m \otimes \mathcal{D}_L^{-1}/(\mathfrak{c}^{-1})$ over $S_B = \mathbb{Z}[[c^{-1+}; \Delta]] \otimes B$. The Tate object carries a canonical μ_N-level:

$$\beta_{can} : \mu_N \otimes \mathcal{D}_L^{-1} \hookrightarrow \mathbb{G}_m \otimes \mathcal{D}_L^{-1} \longrightarrow \mathbf{Tate}_{\mathfrak{c}}(\underline{q}), \tag{6.8}$$

and a canonical differential ω_{can} induced from the differential $dt/t \otimes 1$ on $\mathbb{G}_m \otimes \mathcal{D}_L^{-1}$.

DEFINITION 6.5. Let f be a p-adic Hilbert modular form, $f \in \mathbb{F}(B, k, \mu_N; r)$. The q-expansion of f in the cusp $\mathbf{Tate}_{\mathfrak{c}}(\underline{q})$ is

$$f\left(\mathbf{Tate}_{\mathfrak{c}}(\underline{q}), \omega_{can}, \beta_{can}, \frac{r}{E_\ell(\mathbf{Tate}(q), \omega_{can})}\right) \in \mathbb{Z}[[q^\nu : \nu \in \mathfrak{c}^{-1+}]]^{U_N^2} \otimes_{\mathbb{Z}} B. \tag{6.9}$$

Similarly, for any cusp $(\mathbf{Tate}_{\mathfrak{a},\mathfrak{b},j}(\underline{q}), \beta_N)$ we define the q-expansion of f by

$$f\left(\mathbf{Tate}_{\mathfrak{a},\mathfrak{b},j}(\underline{q}), \omega_{can,j}, \beta_{N,j}, \frac{r}{E_\ell(\mathbf{Tate}_{\mathfrak{a},\mathfrak{b},j}(\underline{q}), \omega_{can})}\right) \in \mathbb{Z}[[q^\nu : \nu \in (\mathfrak{a}\mathfrak{b})^+]]^{U_N^2} \otimes_{\mathbb{Z}} B. \tag{6.10}$$

We call f holomorphic (respectively, cusp) if all its q-expansion lies in $B[[q^\nu : \nu \in (gera\mathfrak{b})^+]]$ for every cusp (respectively, has no constant coefficient at every cusp). We denote the holomorphic (resp. cusp) forms by

$$\mathbb{M}(B, k, \mu_N; r) \quad (\text{resp. } \mathbb{S}(B, k, \mu_N; r)). \tag{6.11}$$

PROPOSITION 6.6. 1. *Take any* $\mathbb{X} \in \{\mathbb{F}, \mathbb{M}, \mathbb{S}\}$. *Then*

$$\mathbb{X}(B, k, \mu_N; r) = \varprojlim \mathbb{X}(B/p^n B, k, \mu_N; r). \tag{6.12}$$

2. *(Koecher's principle) For $g > 1$,*

$$\mathbb{F}(B, k, \mu_N; r) = \mathbb{M}(B, k, \mu_N; r). \tag{6.13}$$

PROOF. All our objects are p-adic. The same proof as in Chapter 4, Proposition 6.6 works.

Koecher's principle follows from the fact that the Tate objects are defined over normal local rings of dimension greater than 1 (See Section 2). Since divisors have codimension 1 we conclude a modular form whose q-expansion is a rational function on the base with no poles outside the closed point, is regular. □

We offer the same remarks as in Page 128.

1. The q-expansion of a p-adic modular form f' induced from a classical modular form f as in (6.5) is the same q-expansion as of f.
2. The q-expansion is injective. See Corollary 6.10.

6.3. The case when p is nilpotent. In this section we let p be a prime and $N \geq 4$ an integer prime to p. As before B be a p-adic ring. We now assume further that p is nilpotent in B. We use our usual notation

$$\mathcal{M}(B, k, \mu_N) \tag{6.14}$$

for classical Hilbert holomorphic modular forms over B, of μ_N-level structure and weight Norm^k. We shall assume that $g \geq 2$. Thus

$$\mathcal{F}(B, k, \mu_N) = \mathcal{M}(B, k, \mu_N) \tag{6.15}$$

where $\mathcal{F}(B, k, \mu_N)$ are the classical modular forms over B, of μ_N-level structure and weight k with possible poles at infinity. As above, $\mathbb{M}(B, k, \mu_N; r)$ is the space of holomorphic p-adic Hilbert modular forms with growth r. Given $j \geq 0$, define a map:

$$\mathcal{M}(B, k + j\ell, \mu_N) \longrightarrow \mathbb{M}(B, k, \mu_N; r) \tag{6.16}$$

$$f \mapsto \widetilde{f}, \quad \widetilde{f}(\underline{A}, \omega, \beta_N, Y)_{/C} = Y^j f(\underline{A}, \omega, \beta_N). \tag{6.17}$$

We claim that $\widetilde{f}$ is a p-adic Hilbert modular form of weight Norm^k. Indeed,

$$\begin{aligned} \widetilde{f}(\underline{A}, \lambda\omega, \beta_N, \mathrm{Norm}\lambda^\ell Y) &= (\mathrm{Norm}\lambda^\ell Y)^j f(\underline{A}, \lambda\omega, \beta_N) \\ &= \mathrm{Norm}\lambda^{j\ell} Y^j \mathrm{Norm}\lambda^{-(k+j\ell)} f(E, \omega, \beta_N) \\ &= \mathrm{Norm}\lambda^{-k} \widetilde{f}(\underline{A}, \omega, \beta_N, Y). \end{aligned} \tag{6.18}$$

Under this map, $E_\ell f$ is sent to $\widetilde{E_\ell f}$, and

$$\begin{aligned} \widetilde{E_\ell f}(\underline{A}, \omega, \beta_N, Y) &= Y^{j+1} f(\underline{A}, \omega, \beta_N) \cdot E_\ell(\underline{A}, \omega, \beta_N) \\ &= r \cdot Y^j f(\underline{A}, \omega, \beta_N) \\ &= r \cdot \widetilde{f}. \end{aligned} \tag{6.19}$$

Therefore, we have obtained a well-defined homomorphism of B-modules as follows:

$$\left(\bigoplus_{j \geq 0} \mathcal{M}(B, k + j\ell, \mu_N) \right) / (E_\ell - r) \longrightarrow \mathbb{M}(B, k, \mu_N; r). \tag{6.20}$$

Here $(E_\ell - r)$ stands for the submodule generated by $\{(E_\ell - r)f : f \in \mathcal{F}(B, k + j\ell, \mu_N)\}$.

The analogue of the isomorphism

$$\left(\bigoplus_{j\geq 0}\mathcal{F}(B,k+j\ell,\mu_N)\right)/(E_\ell - r)\xrightarrow{\cong}\mathbb{F}(B,k,\mu_N;r) \tag{6.21}$$

existing for $g=1$ (Chapter 4, Proposition 6.7) is probably not true. The proof there, recall, was based on the fact that the modular curve with no cusps is affine. However, some information along these lines (e.g. Equation (6.32) has an analogue). Moreover, in analogue to the elliptic case (Chapter 4, Proposition 6.8) we do have

PROPOSITION 6.7. *Let r be a p-adic unit in B any p-adic ring (p not necessarily nilpotent). Then*

$$\mathbb{M}(B,k,\mu_N;r)\cong\left(\oplus_{j=0}^{\infty}\mathcal{M}(B,k+j\ell,\mu_N)\right)/(E_\ell - r), \tag{6.22}$$

and

$$\mathbb{S}(B,k,\mu_N;r)\cong\left(\oplus_{j=0}^{\infty}\mathcal{S}(B,k+j\ell,\mu_N)\right)/(E_\ell - r). \tag{6.23}$$

As in passing from Proposition 6.8 to Theorem 6.9, one deduces the general case from the case where p is nilpotent.

Assume that p is nilpotent. The proof for holomorphic modular forms is practically the same. One needs to use that the ordinary locus $\mathfrak{M}(\mathbb{F},\mu_N)^{\mathrm{ord}}$ (the cusps are ordinary) is affine. This follows from the fact that its complement is (H) and H is a section of an ample line bundle.

For the case of cusps forms one needs that cusp forms are sections of a suitable quasi-coherent sheaf (and thus its higher cohomology on an affine scheme vanishes). This is a bit delicate (and technically demanding). We refer the interested reader to [**38**].

6.4. Katz's expansion. Consider the map:

$$\mathcal{M}(\mathcal{O}_{Lp},k+j\ell,\mu_N)\xrightarrow{\times E_\ell}\mathcal{M}(\mathcal{O}_{Lp},k+(j+1)\ell,\mu_N). \tag{6.24}$$

We note that upon reduction modulo p this map is injective. This implies that the map in (6.24) is injective and splits. We choose complements:

$$\mathcal{M}(\mathcal{O}_{Lp},k+(j+1)\ell,\mu_N)= E_\ell\cdot\mathcal{M}(\mathcal{O}_{Lp},k+j\ell,\mu_N)\ \oplus\ A(\mathcal{O}_{Lp},k+(j+1)\ell,\mu_N). \tag{6.25}$$

($A(\mathcal{O}_{Lp},k,\mu_N)=\mathcal{M}(\mathcal{O}_{Lp},k,\mu_N)$). We may tensor with B and we get the same equality with B-coefficients. Then

$$\bigoplus_{j=0}^{j}A(B,k+a\ell,\mu_N)\cong\mathcal{M}(B,k+j\ell,\mu_N); \tag{6.26}$$

the map given by

$$(f_0,\dots,f_j)\mapsto\sum_{a=0}^{j}f_a\cdot E_\ell^{j-a}. \tag{6.27}$$

Consider the p-adically complete B-module:

$$A^{\mathrm{rigid}}(B,k,\mu_N)= \left\{\sum_{a=0}^{\infty}b_a : b_a\in A(B,k+a\ell,\mu_N),\ b_a\longrightarrow 0\ p\text{-adically uniformly}\right\}. \tag{6.28}$$

By the same methods of Chapter 6.5 we prove:

PROPOSITION 6.8. (Katz's expansion) *For every growth condition r there exists an isomorphism:*

$$A^{\mathrm{rigid}}(B,k,\mu_N) \xrightarrow[\psi]{\sim} \mathbb{M}(B,k,\mu_N;r) \tag{6.29}$$

given by

$$\sum_{a=0}^{\infty} b_a \ \mapsto \ll \sum_{a=0}^{\infty} r^a b_a / E_\ell^a \gg, \tag{6.30}$$

where the right hand side stands for the p-adic Hilbert modular form whose value on a test object $(\underline{A},\omega,\beta_N,Y)_{/A}$ (where $Y\cdot E_\ell(\underline{A},\omega)=r$) is:

$$\ll \sum_{a=0}^{\infty} b_a/E_\ell^a \gg (\underline{A},\omega,\beta_N,Y) = \sum_{a=0}^{\infty} Y^a b_a(\underline{A},\omega,\beta_N). \tag{6.31}$$

6.5. Properties of q-expansions of p-adic modular forms.

PROPOSITION 6.9. *Let $b\in B$ be an element dividing a positive power of p. Let $f\in\mathbb{M}(B,k,\mu_N;1)$. The followings assertions are equivalent:*

1. *$f\in b\cdot\mathbb{M}(B,k,\mu_N;1)$.*
2. *The q-expansion of f lies in $b\cdot B[[q]]$.*

The proof is the same as the proof of Proposition 6.12.

COROLLARY 6.10. *The q-expansion map on p-adic modular forms is injective.*

THEOREM 6.11. *Let $f(q)\in B[[q]]$ be a power series. The following assertions are equivalent:*

1. *$f(q)$ is the q-expansion of an element $f\in\mathbb{M}(B,k,\mu_N;1)$.*
2. *For all n, there exists a positive integer $M(n)$, $M(n)\equiv 0 \mod p^{n-1}$, and a classical modular form $g_n\in\mathcal{M}(B,k+M(n)\ell,\mu_N)$ such that the q-expansion $g_n(q)\equiv f(q) \mod p^n$.*

PROOF. First, let us show that 2 $\implies$ 1. Writing $E_\ell(q)=1+px$, we see that $E_\ell^{p^{n-1}}\equiv 1 \mod p^n$. Now, multiplication of g_n by $E_\ell^{p^{n-1}}$ changes the weight by ℓp^{n-1}, so we can assume $M(n)$ is increasing. Let $\Delta(n)\equiv M(n+1)-M(n)$, so

$$g_{n+1}-g_n\cdot E_\ell^{\Delta(n)} \in p^n\cdot\mathcal{M}(B,k+\ell M(n+1),\mu_N), \tag{6.32}$$

(since $\Delta(n)\equiv 0 \mod p^{n-1}$).

Hence, $g_0+\sum_{a=0}^{\infty}(g_{a+1}-g_a\cdot E_\ell^{\Delta(a)})\in\mathbb{M}(B,k,\mu_N;1)$. Modulo p^n this sum is

$$g_0+(g_1-g_0E_\ell^{\Delta(0}))+\cdots+(g_n-g_{n-1}\cdot E_\ell^{\Delta(n-1)}). \tag{6.33}$$

But $E_\ell=1$ in $\mathbb{M}(B,k,\mu_N;1)$, so the telescopic sum is equal to g_n. Hence, the q-expansion is $\lim g_n(q)=f(q)$.

The implication 1 $\implies$ 2 can be proved as follows: Let $f\in\mathbb{M}(B,k,\mu_N;1)$. Then

$$f=\psi(\sum_{a=0}^{\infty} b_a), \quad b_a\in A(B,k+a\ell,\mu_N). \tag{6.34}$$

Consider $c_n=\psi(\sum_{a=0}^{n} b_a)=\phi(\eta(\sum_{a=0}^{n} b_a))\in\mathcal{M}(B,k+n\ell,\mu_N)$. Take $M(n)$ to be suitably increasing powers of p and $g_n=c_{M(n)}$. □

We didn't define Serre modular forms. But the definition is rather obvious.

DEFINITION 6.12. A p-adic Hilbert modular form (à la Serre) over L_p, of level μ_N, is a uniform p-adic limit of Hilbert modular forms over L, of level μ_N.

Such a modular form would have a limit which is a weight in $\mathbb{Z}_p$. This follows from the results in [**38**].

COROLLARY 6.13. *Serre's p-adic modular forms of weight* $\mathrm{Norm}^k \in \mathbb{Z}$ *are the same as p-adic modular forms à la Katz of growth condition* 1*:* $\mathbb{M}(B, k, \mu_N; 1)$.

CHAPTER 6

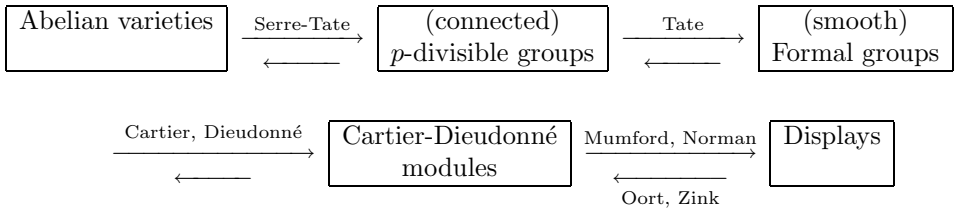
Deformation Theory of Abelian Varieties

The goal of this chapter is to explain the local deformation theory of abelian varieties in characteristic p, concentrating on the case of abelian varieties with real multiplication. Some of the applications we have in mind are proving that the divisor of the Hasse invariant is a reduced normal crossing divisor.

Non-standard conventions is that $\mathbb{N} = \{0, 1, 2, 3, \dots\}$, while $\mathbb{N}^* = \{1, 2, 3, \dots\}$, and that $W(R)$ denotes the "full" Witt vectors while $W_p(R)$ denotes the Witt vectors with components a power of p. The ring $W_p(R)$ is the "usual" ring of Witt vector (e.g., $W_p(\mathbb{Z}/p\mathbb{Z}) = \mathbb{Z}_p$). It is denoted elsewhere in this book by $W(R)$.

Most attributions will be omitted. The results are due foremost to Kodaira-Spencer, Grothendieck, Serre, Tate, Dieudonné, Cartier, Lazard, Mumford, Norman, Deligne, Rapoport, Oort, Zink and others.

Here is a rough diagram of the strategy for studying deformations of abelian varieties we are going to present:



1. Fine moduli schemes

Fix a rational prime p, an integer $g \geq 1$, and an integer $n \geq 3$ such that $(n, p) = 1$. We shall assume we are in one of the following situations :

I. We consider principally polarized abelian schemes $(A, \lambda, \alpha)/_S$ of relative dimension g, and full level-n-structure. Thus, $A \longrightarrow S$ is an abelian scheme of relative dimension g, $\lambda : A \longrightarrow A^t$ is an isomorphism, defined locally on S by a relatively ample line bundle, and $\alpha : (\mathbb{Z}/n\mathbb{Z})^{2g} \longrightarrow A[n]$ is an isomorphism of constant S-group schemes.

There exists a fine moduli space $\mathcal{A} \longrightarrow \mathrm{Spec}(\mathbb{Z})$ for such data. That is, there exists a scheme $\mathcal{A} \longrightarrow \mathrm{Spec}(\mathbb{Z})$ representing the functor

$$S \mapsto \{S - \text{isomorphism classes of } (A, \lambda, \alpha)/_S\}. \tag{1.1}$$

Definition 1.1. Let $S \longrightarrow T$ be a morphism of schemes and $(A, \lambda, \alpha)/_S$ be given. A *deformation* of $(A, \lambda, \alpha)/_S$ to T is a triple $(A', \lambda', \alpha')/_T$ together with an

isomorphism over S

$$(A', \lambda', \alpha') \times_T S \cong (A, \lambda, \alpha). \tag{1.2}$$

Two deformations, (A', λ', α') and $(A'', \lambda'', \alpha'')$, are isomorphic if there exists an isomorphism of abelian schemes over T, say $\varphi : A' \longrightarrow A''$, such that $\varphi|_S$ is the identity, $\varphi^* \lambda'' = \lambda'$, and $\varphi \circ \alpha' = \alpha''$.

Let k be a field. Let $\Lambda = k$ if k has characteristic zero and $\Lambda = W_p(k)$ (infinite Witt vectors) if k has characteristic p. Let $\mathcal{C}_k$ the category of local artinian rings Λ-algebras $(R, \mathfrak{m}_R)$ together with a *given* isomorphism $R/\mathfrak{m}_R \cong k$. Morphisms are local homomorphisms of rings inducing the identity on k.

Fix an object $(A, \lambda, \alpha)/_{\mathrm{Spec}k}$ and consider the functor $\mathcal{C}_k \longrightarrow \underline{\mathrm{Sets}}$ given by

$$R \mapsto \{ \text{ isomorphism classes of deformations over } \mathrm{Spec}(R) \text{ of } (A, \lambda, \alpha)/_k\}. \tag{1.3}$$

This functor is known to be representable by a ring, say R^U, and if $x \in \mathcal{A}(k)$ is the moduli point corresponding to $(A, \lambda, \alpha)/_S$, then

$$R^U \cong \widehat{\mathcal{O}}_{\mathcal{A},x}. \tag{1.4}$$

We shall say we are in the situation I_p if all schemes are $\mathbb{F}_p$ -schemes.

II. We are given a totally real number field L of degree g over $\mathbb{Q}$. We denote by $\mathcal{O}_L$ its ring of integers, by $\mathcal{D}_L$ its different ideal relative to $\mathbb{Q}$ and by d_L its discriminant. We assume that $p \nmid d_L$. Write p as a product of prime ideals

$$p = \mathfrak{p}_1 \cdot \ldots \cdot \mathfrak{p}_r \tag{1.5}$$

in $\mathcal{O}_L$, and let

$$f_i = \deg(\mathfrak{p}_i/p), \quad f = \mathrm{lcm}\{f_i\}. \tag{1.6}$$

Fix a field $\mathbb{F}$ of p^f elements. Note that if we let σ denote the Frobenius automorphism on $\mathbb{F}$ as well as on the Witt vectors $W(\mathbb{F})$, then the embeddings of L in $W(\mathbb{F}) \otimes_{\mathbb{Z}} \mathbb{Q}$ can be indexed by

$$\sigma_{(j,i)}, \quad 1 \leq j \leq r, 1 \leq i \leq f, \tag{1.7}$$

in such a way that

$$\sigma \circ \sigma_{(j,i)} = \sigma_{(j,i+1)}. \tag{1.8}$$

We consider triples $(A, \lambda, \beta_N)/_S$ where S is a $\mathbb{Z}[{d_L}^{-1}]$-scheme, consisting of an abelian variety $A \longrightarrow S$ of relative dimension g, an embedding $\iota : \mathcal{O}_L \longrightarrow \mathrm{End}(A/S)$ such that condition (**R**) holds (Chapter 3, Section 5), and an $\mathcal{O}_L$-equivariant embedding $\beta_N : \mu_N \otimes \mathcal{D}_L^{-1} \longrightarrow A$.

There exists a fine moduli scheme $\mathfrak{M}(\mu_N) \longrightarrow \mathbb{Z}[d_L^{-1}]$ for such data. The same remarks concerning deformations, deformations over $\mathcal{C}_k$ etc. apply here with the obvious modifications. We shall say we are in case II_p if all schemes are $\mathbb{F}$-schemes.

Most local deformations theories over p-adic rings for I or II (or similar scenarios) pass through the *p-divisible group*,

$$A(p) = \varinjlim_n A[p^n], \tag{1.9}$$

of the abelian variety A. See [**11**] for a review of the available tools and [**113**], [**104**], for essentials on p-divisible groups.

From now on, let k stand for a fixed field of characteristic p. Let A/k be an abelian variety.

THEOREM 1.2. (Serre-Tate) *For every ring R in $\mathcal{C}_k$ the functor*

$$\mathbb{A} \mapsto \mathbb{A}(p), \tag{1.10}$$

induces an equivalence of categories between the category of deformations $\mathbb{A}$ of A over R with morphisms being morphisms of abelian schemes over R, to the category of deformations of $A(p)$ into p-divisible groups over R with the morphisms being morphisms of p divisible groups whose restriction to $A(p)$ comes from an endomorphism of A/k.

One immediately deduces that in case I we have an equivalence of categories

$$\left\{\begin{array}{c}\text{Isomorphism classes of}\\ \text{deformations of } (A,\lambda,\alpha)/k\\ \text{to } R\end{array}\right\} \longrightarrow \left\{\begin{array}{c}\text{Isomorphism classes of}\\ \text{deformations of } (A(p),\lambda(p))/k\\ \text{to } R\end{array}\right\}, \tag{1.11}$$

where $\lambda(p) : A(p) \longrightarrow A^t(p) \cong (A(p))^t$ is deformed into symmetric isomorphism $\lambda : G \longrightarrow G^t$.

In case II we have an equivalence of categories

$$\left\{\begin{array}{c}\text{Isomorphism classes of}\\ \text{deformations of } (A,\iota,\alpha)/k\\ \text{to } R\end{array}\right\} \longrightarrow \left\{\begin{array}{c}\text{Isomorphism classes of}\\ \text{deformations of } (A(p),\iota(p))/k\\ \text{to } R\end{array}\right\}, \tag{1.12}$$

where $\iota(p) : \mathcal{O}_L \longrightarrow \mathrm{End}(A(p))$ is deformed to homomorphisms $\iota : \mathcal{O}_L \longrightarrow \mathrm{End}(G)$.

REMARK 1.3. Note that the extra level structure disappears. This follows from rigidity of étale group schemes; i.e., there always exists a way to extend this level structure to any such deformation, and uniquely so. Note also that we do not require condition (**R**) to hold for a deformation. It is automatically satisfied.

EXAMPLE 1.4. Let E/k be an ordinary elliptic curve over k. Then

$$E(p) \cong \underline{\mathbb{Q}_p/\mathbb{Z}_p}_{/k} \oplus \widehat{\mathbb{G}_{m/k}}, \tag{1.13}$$

where

$$\widehat{\mathbb{G}_m} = \varinjlim \ \mu_{p^n}. \tag{1.14}$$

We find a coarser formulation of the Serre-Tate coordinates discussed in Chapter 3, Section 4.2; there is a canonical deformation, or *canonical lift*, of $E(p)$ to R. Namely

$$\underline{\mathbb{Q}_p/\mathbb{Z}_p}_{/R} \oplus \widehat{\mathbb{G}_{m/R}}. \tag{1.15}$$

Clearly it has the property that every endomorphism of $E(p)$ extends to it. We conclude that for every $R \in \mathcal{C}_k$ there exists a canonical lift of E over R, say E^{can}/R with the property

$$\mathrm{End}_R(E^{can}) \cong \mathrm{End}_k(E). \tag{1.16}$$

The same argument applies to ordinary abelian varieties.

2. Proof of the Serre-Tate theorem

We follow Drinfeld's proof as presented in [**?**]. Consider a more general scenario: We are given two rings, R and R_0, and a ring homomorphism $\varphi : R \longrightarrow R_0$ with kernel I, such that there exists a b with $I^b = 0$ and there exists an a such that $p^a = 0$ in R_0. We consider two categories:

• $\underline{\mathrm{Ab}}_R$ = Category of abelian schemes over R with morphisms being homomorphisms of abelian schemes.

• $\underline{\mathrm{Def}}_R$ = Category of triples $(A/R_0, G/R, \iota)$, where A/R_0 is an abelian scheme, G/R is a p-divisible group and $\iota : A(p) \longrightarrow G|_{R_0}$ is an isomorphism. An arrow $(A/R_0, G/R, \iota) \longrightarrow (A'/R_0, G'/R, \iota')$ is a homomorphism $f : G \longrightarrow G'$ whose restriction $f_0 : A(p) \longrightarrow A'(p)$, induced via ι and ι', is induced also from a homomorphism of abelian schemes $A \longrightarrow A'$.

We have a natural functor

$$\Phi : \underline{\mathrm{Ab}}_R \longrightarrow \underline{\mathrm{Def}}_R; \quad \Phi(A) = (A \otimes_R R_0, A(p)/R, \iota_{can}). \tag{2.1}$$

THEOREM 2.1. (Serre-Tate) *The functor Φ is an equivalence of categories.*

PROOF. We first make a

DEFINITION 2.2. Let G be either a finite flat group scheme over R or an abelian scheme over R. Let J be an ideal of R. Define a group functor G_J on $\underline{R - \mathrm{Alg}}$ by

$$G_J(S) = \mathrm{Ker}(G(S) \longrightarrow G(S/J)), \tag{2.2}$$

where here and below S/J stands for $S \otimes_R (R/J)$.

LEMMA 2.3. *Let $q = p^{ab}$. Then*

$$[q]G_I = 0. \tag{2.3}$$

PROOF. We assume that G is an abelian scheme. After localizing on R, we may assume that $G = \mathrm{Spf} R[[x_1, \ldots, x_d]]$ with augmentation ideal topologically generated by $x_1, \ldots, x_d$ (for a finite flat group scheme, write $G = \mathrm{Spec}\, R[x_1, \ldots, x_d]/\mathfrak{a}$ for a suitable ideal $\mathfrak{a}$ and augmentation idea generated by $x_1, \ldots, x_d$).

Let $S \in \underline{R - \mathrm{Alg}}$ and $\alpha \in G_J(S)$. Then we may identify α with a vector $(\alpha_1, \ldots, \alpha_d)$ with $\alpha_i \in JS$. Writing

$$([p^a]\alpha)_i = p^a \alpha_i + \text{higher order terms in } \alpha_1, \ldots, \alpha_d, \tag{2.4}$$

we see that $[p^a]\alpha \in G_{IJ+J^2}$. Applying this to $J = I, I^2, I^3, \ldots$ we see that $[q]G_I = 0$. □

Note that the lemma also holds for a p-divisible group over R.

PROPOSITION 2.4. *Let G and H be either p-divisible groups or abelian schemes over R. Let G_0 and H_0 be the objects obtained from G and H, respectively, by base change to R_0.*

1. The groups $\mathrm{Hom}_R(G, H)$ and $\mathrm{Hom}_{R_0}(G_0, H_0)$ are torsion free.

2. The homomorphism $\mathrm{Hom}_R(G, H) \longrightarrow \mathrm{Hom}_{R_0}(G_0, H_0)$ is injective.

3. Given f_0 in $\mathrm{Hom}_{R_0}(G_0, H_0)$ there exists a unique homomorphism $\widetilde{qf_0}$ in $\mathrm{Hom}_R(G, H)$ lifting qf_0.

4. The homomorphism f_0 lifts to a homomorphism $\widetilde{f_0}$ in $\mathrm{Hom}_R(G, H)$ if and only if $G[q] \subset \mathrm{Ker}(\widetilde{qf_0})$.

PROOF. Part 1 is clear. For 2, we note that

$$\mathrm{Ker}(\mathrm{Hom}_R(G,H) \longrightarrow \mathrm{Hom}_{R_0}(G_0,H_0)) = \mathrm{Hom}_R(G,H_I). \tag{2.5}$$

But H_I is torsion and $\mathrm{Hom}_R(G,H)$ is torsion free!

3. Let $\alpha \in G(S)$ and let $\bar{\alpha}$ denote its image in $G(S/I)$. Regarding $\bar{\alpha}$ as an element of $G_0(S/I)$, we write $f_0(\bar{\alpha})$ for the image in $H_0(S/I)$. Since the map $H(S) \longrightarrow H(S/I)$ is surjective, we may lift $f_0(\bar{\alpha})$ to some $\alpha' \in H(S)$. Let

$$\widetilde{qf_0}(\alpha) = q\alpha'. \tag{2.6}$$

This is well defined, because any other choice of a lift, say α'' differs from α' by an element of $H_I(S)$, hence killed by q. Immediate verification shows that $\widetilde{qf_0}$ is a homomorphism and its uniqueness follows from part 2.

4. Clearly if f_0 lifts to some $\widetilde{f_0}$ then, by uniqueness, $\widetilde{qf_0} = q\widetilde{f_0}$ and therefore $\widetilde{qf_0}$ kills $G[q]$. Conversely, if $\widetilde{qf_0}$ kills $G[q]$ we may factor it as $\widetilde{qf_0} = qg$ for some $g \in \mathrm{Hom}_R(G,H)$. Let $g_0 \in \mathrm{Hom}_{R_0}(G_0,H_0)$ be the homomorphism induced by g. Then $qg_0 = qf_0$ as both are obtained by restricting $\widetilde{qf_0}$. Since $\mathrm{Hom}_{R_0}(G_0,H_0)$ is torsion free, $g_0 = f_0$. □

Let us show now that Φ is fully-faithful. I.e., that

$$\mathrm{Hom}_R(A,B) \longrightarrow \mathrm{Hom}((A_0,A(p)),(B_0,B(p))) \tag{2.7}$$

is an isomorphism. We know it is injective by Proposition 2.4. Let, therefore, f be in $\mathrm{Hom}(A_0,B_0)$ and ϕ be in $\mathrm{Hom}(A(p),B(p))$ such that

$$f(p) = \phi_0, \tag{2.8}$$

where here and below "(p)" denotes passing to the p-divisible and the subscript "$_0$" denotes base change to R_0. Using Proposition 2.4, we see that the homomorphism qf lifts uniquely to a homomorphism $\widetilde{qf} \in \mathrm{Hom}_R(A,B)$ and f lifts to $\mathrm{Hom}_R(A,B)$ if and only if $\widetilde{qf}$ kills $A[q]$. But $\widetilde{qf}$ kills $A[q]$ if and only if $\widetilde{qf}(p)$ kills $A[q]$. Since $\widetilde{qf}(p)$ is a lift of $q\phi_0$ to $A(p)$ and since ϕ_0 *does* lift to $A(p)$, we have, by part 4 of Proposition 2.4 that $\widetilde{qf}(p)$ kills $A[q]$. Therefore f lifts.

It remains to prove essential surjectivity of Φ. Let (A_0,G,ι) be an object of $\underline{\mathrm{Def}}_R$. Because $R \longrightarrow R_0$ is a nilpotent thickening there exists *some* abelian scheme B lifting A_0 to R. This can be deduced for example from [**83**, 6.3]. Let α_0 be the isomorphism $B(p)_0 \longrightarrow G_0$. Let $\widetilde{q\alpha_0} \in \mathrm{Hom}_R(B(p),G)$ be the canonical lift of $q\alpha_0$. Let $K = \mathrm{Ker}(\widetilde{q\alpha_0})$. Then K is a finite flat group scheme such that $K_0 = A_0[q]$, and $D := B/K$ is an abelian scheme. The map $\widetilde{q\alpha_0}$ induces an isomorphism between $D(p)$ and G, which under the identification $D_0 = A_0/A_0[q]$ is equal to α_0. □

3. Deformation of p-divisible groups

Let R be a ring. By an n-dimensional smooth commutative formal group over R, or simply, *n-dimensional formal group*, we mean the following: It is a power series ring $R[[x_1,\ldots,x_n]]$ together with n power series

$$F(\underline{x},\underline{y}) = (F_1(\underline{x},\underline{y}),\ldots,F_n(\underline{x},\underline{y})) \tag{3.1}$$

in the $2n$ variables $\underline{x} = (x_1,\ldots,x_n)$ and $\underline{y} = (y_1,\ldots,y_n)$. The following are required to hold for every i:

- $F_i(\underline{x},0) = x_i$, $F_i(0,\underline{y}) = y_i$.

- $F_i(\underline{x},\underline{y}) = F_i(\underline{y},\underline{x})$.
- $F_i(F(\underline{x},\underline{y}),\underline{z}) = F_i(\underline{x},F(\underline{y},\underline{z}))$.

Define inductively the multiplication by m maps $[m]$:

$$[1](\underline{x}) = \underline{x};\ \ [m](\underline{x}) = F([m-1](\underline{x}),\underline{x}). \tag{3.2}$$

Note that $[m](\underline{x}) = ([m]_1(\underline{x}),\dots,[m]_n(\underline{x}))$ defines a endomorphism of $R[[\underline{x}]]$ by

$$x_i \mapsto [m]_i(\underline{x}). \tag{3.3}$$

We say that the formal group is *p-divisible* if $R[[x]]$, considered as an $R[[x]]$ module via $[p](\underline{x})$, is a finitely generated module.

EXAMPLE 3.1. Let $n = 1$. Define the *additive formal group* $\widehat{\mathbb{G}_{a\ /R}}$ over any ring R by

$$F_a(x,y) = x + y. \tag{3.4}$$

Thus,

$$[p](x) = px. \tag{3.5}$$

One verifies that $\widehat{\mathbb{G}_a}$ is p-divisible if and only if $p \in R^\times$. Remark that $\widehat{\mathbb{G}_a}/R$ is the completion of $\mathbb{G}_a/R = \mathrm{Spec}(R[x])$ along the zero section defined by the ideal sheaf (x).

EXAMPLE 3.2. Let $n = 1$. Define the *multiplicative formal group* $\widehat{\mathbb{G}_{m\ /R}}$ over any ring R by

$$F_m(x,y) = x + y - xy. \tag{3.6}$$

Equivalently, $1 - F_m(x,y) = (1-x)(1-y)$. Thus,

$$1 - [p](x) = (1-x)^p. \tag{3.7}$$

One verifies that $\widehat{\mathbb{G}_m}$ is p-divisible for every p. In fact $\{1, x, \dots, x^{p-1}\}$ span $R[[x]]$ as a module (via $[p](x)$) over $R[[x]]$. Remark that $\widehat{\mathbb{G}_m}/R$ is the completion of $\mathbb{G}_m/R = \mathrm{Spec}(R[t,t^{-1}]) = \mathrm{Spec}(R[1-x,(1-x)^{-1}])$, where $x = 1-t$, along the zero section defined by the ideal sheaf (x).

Note that F_m is the group law of $1 - \exp(-x)$. That is

$$1 - \exp(-(x+y)) = F_m(1-\exp(-x), 1-\exp(-y)). \tag{3.8}$$

EXAMPLE 3.3. Let ϵ and δ be free variables. Define a power series $F_{\epsilon,\delta}(x,y)$ in $\mathbb{Z}[1/2][[\epsilon,\delta]]$ by

$$F_{\epsilon,\delta}(x,y) = \frac{x\sqrt{\ell(y)} + y\sqrt{\ell(x)}}{1 - \epsilon x^2 y^2},\ \ \ell(x) = 1 - 2\delta x^2 + \epsilon x^4. \tag{3.9}$$

The fact that this is a formal group law is quite tedious to check directly. A more insightful approach is to prove that this is the formal group associated to the elliptic curve $E:\ y^2 = \ell(x)$ with the zero point $(0,1)$.

As a special case, take ϵ and δ equal to one. Then the formal group is simply

$$F(x,y) = \frac{x+y}{1+xy}. \tag{3.10}$$

It is immediate to verify the axioms in this case. Note that $F(x,y)$ is the group law of the hyperbolic tangent

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}. \tag{3.11}$$

That is,

$$\tanh(x+y) = F(\tanh(x), \tanh(y)). \tag{3.12}$$

THEOREM 3.4. (Tate [**113**]) *Let $(R, \mathfrak{m})$ be a complete noetherian local ring with residue field $k = R/\mathfrak{m}$ a field of characteristic p. Then the category of connected p-divisible over R is equivalent to the category of p-divisible smooth commutative formal groups over R.*

Note that every R in $\mathcal{C}_k$ is of this form. Thus, to study local deformations of data I or II, one is led to consider p-divisible formal groups over local artinian rings. We remark that the situation that will ultimately interest us is when both the p-divisible group and its dual are connected. To deal with the general situation one uses the theory of mixed extensions (see [**11**]). We further remark that certain strengthening of Tate's theorem are known. For example, if one generalizes the notion of a formal group to objects "looking locally like a formal group", then Tate's theorem holds for p-adic rings R (i.e., $R \cong \varprojlim R/p^n R$) – see [**69**].

PROOF. **To be supplied.** □

4. Commutative smooth formal groups

We follow Lazard [**Laz**] in our exposition of the study of commutative smooth formal groups, or simply, formal groups, via their Cartier-Dieudonné modules. We shall refer for most details to loc. cit..

Before embarking, we remark that while Cartier-Dieudonné theory works for *any* formal group (commutative, smooth) over *any* commutative associative ring with 1, the *applications* to studying deformations of p-divisible groups are limited to specific rings, e.g., artinian local rings.

Let K be any commutative associative ring with 1 and denote by $\underline{\mathrm{nil}(K)}$ the category of commutative nil-algebras A over K. This means that every element in A is nilpotent. Note that A has no identity element. We also let $\underline{\mathrm{nil}(K,n)}$ be the full sub-category of $\underline{\mathrm{nil}(K)}$ consisting of nil-algebras A such that for every $x_1, \ldots, x_{n+1}$ in A we have $x_1 \cdot \ldots \cdot x_{n+1} = 0$.

We are interested in three closely related functors from the category $\underline{\mathrm{nil}(K)}$ to the category of pointed sets $\underline{*\,\mathrm{Sets}}$. They are the following:

- $\underline{\mathrm{Models}}$: These are the functors

$$D^{(I)} : \underline{\mathrm{nil}(K)} \longrightarrow \underline{*\,\mathrm{Sets}}; \quad D^{(I)}(A) = A^{(I)}, \tag{4.1}$$

where I is any index set and $A^{(I)} = \{(a_\alpha)_{\alpha \in I} : a_\alpha = 0 \text{ except for finitely many } \alpha\}$.

- $\underline{\mathrm{Formal\ modules}}$: For any *free* K-module L, define a functor

$$L^+ : \underline{\mathrm{nil}(K)} \longrightarrow \underline{*\,\mathrm{Sets}}; \quad L^+(A) = A \otimes_K L. \tag{4.2}$$

• Formal varieties: A formal variety over K is a functor

$$\mathcal{V} : \underline{\mathrm{nil}(K)} \longrightarrow \underline{*\,\mathrm{Sets}} \tag{4.3}$$

isomorphic to $D^{(I)}$ for some I.

Thus, formal modules are examples of formal varieties. A choice of a K-basis provides such an isomorphism. Formal varieties appear naturally in algebraic geometry in the following way:

Let V be an n-dimensional algebraic variety over K and let $v \in V$ be a non-singular K-rational point. The completion of the local ring of v is isomorphic, by a choice of local coordinates, to $K[[x_1, \ldots, x_n]]$. The functor defined on $\underline{\mathrm{nil}(K)}$ by

$$A \mapsto \mathrm{Hom}_K^{\mathrm{Cont}}(K[[x_1, \ldots, x_n]], K \oplus A), \tag{4.4}$$

where

$$\mathrm{Hom}_K^{\mathrm{Cont}}(K[[x_1, \ldots, x_n]], K \oplus A) := \varprojlim \mathrm{Hom}_k(K[[x_1, \ldots, x_n]]/I^j, K \oplus A) \tag{4.5}$$

and I is the maximal ideal of $K[[x_1, \ldots, x_n]]$, is identified with D^n. More canonically, the completion of V along v, $\widehat{V}$, defines a formal variety taking every nil-K-algebra A to $\mathrm{Mor}(\mathrm{Spec}(K \oplus A), \widehat{V})$ (morphisms in the category of formal varieties).

A particular case of great importance is when V is an abelian variety (resp. scheme) and v is its zero point (resp. section). This includes Example 3.3 as a special case. The case when V is the additive group (resp. multiplicative group) leads to Example 3.1 (resp. Example 3.2).

Formal varieties form a category whose morphisms are functorial morphisms (natural transformation) of functors. If $\mathcal{V}$ and $\mathcal{W}$ are formal varieties, we denote by $\mathrm{Mor}(\mathcal{V}, \mathcal{W})$ the morphisms from $\mathcal{V}$ to $\mathcal{W}$.

LEMMA 4.1. (Morphism Lemma, [**Laz**, p.10]) *The functorial morphisms*

$$f : D^{(I)} \longrightarrow D, \tag{4.6}$$

are in bijection with formal power series $\sum_{\alpha \in \mathbb{N}^{(I)}} c_\alpha x^\alpha$ *with* $c_0 = 0$ *(multi index notation). A morphism* $f : D^{(I)} \longrightarrow D^{(J)}$ *is defined by its components* $f_j : D^{(I)} \longrightarrow D$, $j \in J$.

PROOF. The only point here is that any morphism $f : D^{(I)} \longrightarrow D$ can be expressed as a power series. Given $\lambda = (J, n)$ with J a finite subset of I and n an integer, let A_λ be freely generated in $\underline{\mathrm{nil}(K, n)}$ by elements $\{x_{\lambda,i} : i \in J\}$.

Consider the point x_λ in $D^{(I)}(A_\lambda)$ given by $x_\lambda = (x_{\lambda,i})_{i \in I}$, where $x_{\lambda,i} = 0$ for $i \notin J$. Then $f_{A_\lambda}(x_\lambda)$ is an element of A_λ and can thus be written as

$$f_{A_\lambda}(x_\lambda) = \sum_{\substack{\alpha \in \mathbb{N}^{(I)},\, 0 < |\alpha| \leq n \\ \mathrm{supp}(\alpha) \subset J}} c_{\lambda,\alpha} x^\alpha \tag{4.7}$$

for uniquely determined $c_{\lambda,\alpha} \in K$. But for any $A \in \underline{\mathrm{nil}(K)}$ and any $x \in A$ we can find $\lambda = (J, n)$ and $\phi : A_\lambda \longrightarrow A$ such that $x = \phi(x_\lambda)$. By functoriality

$$f_A(x) = \sum_{\substack{\alpha \in \mathbb{N}^{(I)},\, 0 < |\alpha| \leq n \\ \mathrm{supp}(\alpha) \subset J}} c_{\lambda,\alpha} x^\alpha. \tag{4.8}$$

Finally, applying this consideration for the case $x = x_{\lambda'} \in A_{\lambda'}$ for $\lambda' = (J', n')$ and $\lambda = (J, n)$ such that $J' \subset J$ and $n' \leq n$, one sees that $c_{\lambda,\alpha}$ depend only on α. □

4.1. Curves. Let $\mathcal{V}$ be a formal variety. A *curve* γ in $\mathcal{V}$ is a morphism

$$\gamma : D \longrightarrow \mathcal{V}. \tag{4.9}$$

Put

$$\mathcal{C}(\mathcal{V}) = \mathrm{Mor}(D, \mathcal{V}) \tag{4.10}$$

– the set of curves in $\mathcal{V}$. One defines $\mathcal{T}\gamma$ to be the induced map

$$\mathcal{T}\gamma : D|_{\underline{\mathrm{nil}(K,1)}} \longrightarrow \mathcal{V}|_{\underline{\mathrm{nil}(K,1)}}, \tag{4.11}$$

and calls it the *tangent* to γ. We further let

$$\mathcal{T}\mathcal{V} = \{\mathcal{T}\gamma : \gamma \in \mathcal{C}(\mathcal{V})\}, \tag{4.12}$$

and call it the *tangent space* to $\mathcal{V}$. We make $\mathcal{T}$ into a functor by defining $\mathcal{T}f$ for $f : \mathcal{V} \longrightarrow \mathcal{W}$ by the formula

$$\mathcal{T}f \circ \mathcal{T}\gamma = \mathcal{T}(f \circ \gamma). \tag{4.13}$$

One easily verifies, using the Morphism Lemma, that any curve in $\mathcal{C}(D^{(I)})$ is of the form $\gamma = \sum_{n\geq 1} a_n t^n$ for some $a_n \in K^{(I)}$ and vice versa. Then $\mathcal{T}\gamma$ is just $a_1 \in K^{(I)}$. Beside showing that the definition of tangent agrees with the intuitive one, it also gives a canonical isomorphism of $\mathcal{T}D^{(I)}$ with $K^{(I)}$, and thus $\mathcal{T}D^{(I)}$ is canonically a free K-module. One can show further that given a formal variety $\mathcal{V}$, the K-linear structure on $\mathcal{T}\mathcal{V}$ deduced from a choice of isomorphism $\mathcal{V} \cong D^{(I)}$ is in fact independent of the choice of isomorphism. Furthermore, if $f : \mathcal{V} \longrightarrow \mathcal{W}$ is a morphism then $\mathcal{T}f$ is K-linear. One puts:

$$\dim(\mathcal{V}) = \dim_K(\mathcal{T}\mathcal{V}). \tag{4.14}$$

The concept of curves is crucial to the whole theory we are about to present. The following results are evidence for that.

Theorem 4.2. (Isomorphism Theorem, [**Laz**, I.8]) *A morphism of formal varieties $f : \mathcal{V} \longrightarrow \mathcal{W}$ is an isomorphism if and only if $\mathcal{T}f : \mathcal{T}\mathcal{V} \longrightarrow \mathcal{T}\mathcal{W}$ is an isomorphism.*

The proof is quite easy. One direction is immediate. For the other, one passes to models and applies the Morphism Lemma, which reduces the assertion to *formal* inversion of power series.

Lemma 4.3. (Curves Lemma, [**Laz**, I.10]) *The functor $\mathcal{V} \mapsto \mathcal{C}(\mathcal{V})$ is faithful. That is, if f and f' are morphisms from $\mathcal{V}$ to $\mathcal{W}$ and $f \circ \gamma = f' \circ \gamma$ for every $\gamma \in \mathcal{C}(\mathcal{V})$, then $f = f'$.*

The argument is easy: Reduce to the the case $\mathcal{V} = D^{(I)}$ and $\mathcal{W} = D$ and use the morphism lemma and enough "test curves" to show that $f - f'$ (the difference defined using the power series expression) is actually zero.

4.2. Formal groups. A formal variety $\mathcal{G}$ is called a *formal group* if $\mathcal{G}$ is a commutative group object in the category of formal varieties. This amounts to giving morphisms

$$m : \mathcal{G} \times \mathcal{G} \longrightarrow \mathcal{G}, \quad inv : \mathcal{G} \longrightarrow \mathcal{G}, \tag{4.15}$$

such that the expected diagrams commute. Still more concretely, it is equivalent to giving for every $A \in \underline{\mathrm{nil}(K)}$ a group structure on $\mathcal{G}(A)$ that is functorial in A.

We see that if $\mathcal{G} = D^{(I)}$ this agrees with the previous notion of (smooth commutative) formal groups defined in Section 3 using power series. Note that the coordinates of the formal power series $F(\underline{x}, \underline{y}) = (F_1(\underline{x}, \underline{y}), \ldots, F_n(\underline{x}, \underline{y}))$ are the coordinates for the morphism $m : D^n \times D^n \longrightarrow D^n$ constructed from F. The simplest examples of formal groups are thus $\widehat{\mathbb{G}_a} = D$ with $F(x, y) = x + y$ and $\widehat{\mathbb{G}_m} = D$ with $F(x, y) = x + y - xy$.

A more interesting example is provided by the completion $\widehat{A}$ of an abelian variety A at its identity. We explained above that $\widehat{A}$ is a formal variety, and we note that the multiplication morphism $m : A \times A \longrightarrow A$ induces a multiplication map

$$\widehat{A} \times \widehat{A} \longrightarrow \widehat{A}, \tag{4.16}$$

making it into a formal group. See also Example 3.3.

If $\mathcal{G}$ is a formal group then $\mathcal{V} \mapsto \mathrm{Mor}(\mathcal{V}, \mathcal{G})$ is a contravariant functor from formal varieties to abelian groups. In particular $\mathcal{C}(\mathcal{G})$ is an abelian group! We shall write the addition in this group as

$$\gamma_1 \underset{\mathcal{G}}{+} \gamma_2. \tag{4.17}$$

4.3. Operators on $\mathcal{C}(\mathcal{G})$. Given $\varphi \in \mathrm{Mor}(D, D)$ we define a composition operator,

$$\mathrm{comp}(\varphi) : \mathcal{C}(\mathcal{G}) \longrightarrow \mathcal{C}(\mathcal{G}), \tag{4.18}$$

by

$$\mathrm{comp}(\varphi) \cdot \gamma = \gamma \circ \varphi. \tag{4.19}$$

Note that

$$\mathrm{comp}(\varphi_1 \circ \varphi_2) = \mathrm{comp}(\varphi_2) \circ \mathrm{comp}(\varphi_1). \tag{4.20}$$

We define the following composition operators:

- For $n \geq 1$, let V_n denote $\mathrm{comp}(\varphi)$ for $\varphi(t) = t^n$.
- For $c \in K$, let $[c]$ denote $\mathrm{comp}(\varphi)$ for $\varphi(t) = ct$.
- For $n \geq 1$, define an operator F_n as follows: Fix an isomorphism $\mathcal{G} \cong D^{(I)}$. Given $\gamma \in \mathcal{C}(\mathcal{G})$, let $\sigma_\gamma : D^n \longrightarrow \mathcal{G}$ be defined by

$$\sigma_\gamma(t_1, \ldots, t_n) = \gamma(t_1) \underset{\mathcal{G}}{+} \ldots \underset{\mathcal{G}}{+} \gamma(t_n). \tag{4.21}$$

One can prove ([**Laz**, I.11]) that σ_γ factors as $s_\gamma \circ \mathrm{sym}_n$, where

$$\mathrm{sym}_n(x_1, \ldots, x_n) = (\sum x_i, \sum_{i<j} x_i x_j, \ldots, x_1 \cdot \ldots \cdot x_n). \tag{4.22}$$

One defines

$$F_n \cdot \gamma(t) = s_\gamma(0, \ldots, 0, (-1)^{n-1} t). \tag{4.23}$$

Consider $\mathcal{C}$ as a functor on the category of formal groups, and consider the ring of natural transformations of this functor. Denote it by $\mathrm{Cart}(K)$. Every composition operator we defined above is an element of $\mathrm{Cart}(K)$.

THEOREM 4.4. *In* $\mathrm{Cart}(K)$ *the following relations hold:*

$$[1_K] = V_1 = F_1 = 1_{\mathrm{Cart}(K)};\ \ [c][d] = [cd];\ \ F_nF_m = F_{nm};\ \ V_nV_m = V_{nm}; \tag{4.24}$$

$$[c]V_n = V_n[c^n];\ \ F_n[c] = [c^n]F_n;\ \ F_nV_n = n \cdot 1_{\mathrm{Cart}(K)}; \tag{4.25}$$

$$V_nF_m = F_mV_n, \ \ \textit{if}\ \ (n,m) = 1. \tag{4.26}$$

Most of the properties are easy to verify from the definition. For the proof see [**Laz**, I.10, III.3, IV.1]. Note that in general $[c] + [d] \neq [c+d]$, and in particular $n \cdot 1_{\mathrm{Cart}(K)} \neq [n]$. Also, in general, $V_nF_n \neq n \cdot 1_{\mathrm{Cart}(K)}$.

EXAMPLE 4.5. The formal group $\widehat{\mathbb{G}_a}$ (compare Example 3.1):
Let $\gamma(t) = \sum a_n t^n$ and $\gamma' = \sum a'_n t^n$ be two elements of $\mathcal{C}(\widehat{\mathbb{G}_a})$. Then

$$\gamma \underset{\widehat{\mathbb{G}_a}}{+} \gamma'(t) = \sum (a_n + a'_n)t^n, \tag{4.27}$$

and

$$[c]\gamma(t) = \sum a_n c^n t^n,\ \ V_m\gamma(t) = \sum a_n t^{mn},\ \ F_m\gamma(t) = \sum m a_{mn} t^n. \tag{4.28}$$

In particular, letting $\gamma_{\underline{a}}(t) = t$, we get $F_m\gamma_{\underline{a}} = 0$ for all $m > 1$. See [**Laz**, III.3].

EXAMPLE 4.6. The formal group $\widehat{\mathbb{G}_m}$ (compare Example 3.2):
Let γ and γ' be as above. Then

$$\gamma \underset{\widehat{\mathbb{G}_m}}{+} \gamma'(t) = \sum a''_n t^n, \tag{4.29}$$

where

$$a''_n = a_n + a'_n - \sum_{0<i<n} a_i a'_{n-i}. \tag{4.30}$$

We also have

$$[c]\gamma(t) = \sum a_n c^n t^n,\ \ V_m\gamma(t) = \sum a_n t^{mn}. \tag{4.31}$$

Let $\gamma_{\underline{m}}(t) = t$, then $F_n\gamma_{\underline{m}} = \gamma_{\underline{m}}$ for all n.

How does one compute F_n? There is a formal "trick" for that. Let $\mathcal{G}$ be a formal group. One has formally the identity

$$V_mF_m\gamma = \sum_{i=1}^{m}{}_{\mathcal{G}} [\zeta^i]\gamma, \tag{4.32}$$

where ζ is a primitive m-th root of unity (which need not exist in K) and where $\sum_{\mathcal{G}}$ denotes summation $\underset{\mathcal{G}}{+}$. Thus, for example, $V_mF_m\gamma_{\underline{m}} = \zeta t \underset{\widehat{\mathbb{G}_m}}{+} \ldots \underset{\widehat{\mathbb{G}_m}}{+} \zeta^m t =$

$1-\prod_{i=1}^{m}(1-\zeta^i t)=t^m$, which gives $F_m\gamma_{\underline{m}}=\gamma_{\underline{m}}$. More generally,

$$\begin{aligned} V_mF_m\gamma &= (\sum a_n\zeta^n t^n) \underset{\widehat{\mathbb{G}_m}}{+} \dots \underset{\widehat{\mathbb{G}_m}}{+} (\sum a_n\zeta^{mn}t^n) \\ &= 1-\prod_{i=1}^{m}(1-\sum_n a_n\zeta^{in}t^n). \end{aligned}$$

DEFINITION 4.7. Let $\mathcal{V}$ be a formal variety and let $(\gamma_i)_{i\in I}$ be an indexed set of curves in $\mathcal{V}$. It is a called a *basic set of curves* if $(\mathcal{T}\gamma_i)_{i\in I}$ is a K-basis for $\mathcal{T}\mathcal{V}$.

PROPOSITION 4.8. ([**Laz**, III 6.1]) *Let $\mathcal{G}$ be a formal group and $(\gamma_i)_{i\in I}$ be a basic set of curves. Then any curve $\gamma\in\mathcal{C}(\mathcal{G})$ can be written uniquely as*

$$\gamma=\sum_{m\geq 1,i\in I}{}_{\mathcal{G}}\, V_m[x_{m,i}]\gamma_i,\quad \forall m\ (x_{m,i})_{i\in I}\in K^{(I)}. \tag{4.33}$$

Conversely, every such expression defines a curve $\gamma\in\mathcal{C}(\mathcal{G})$.

EXAMPLE 4.9. In the formal group $\widehat{\mathbb{G}_a}$, the set $\{\gamma_{\underline{a}}\}$ is a basic set of curves. Every curve can be written uniquely as

$$\sum_{n=1}^{\infty}a_nt^n=\sum_{n=1}^{\infty}{}_{\mathcal{G}}\, V_n[a_n]\gamma_{\underline{a}}. \tag{4.34}$$

EXAMPLE 4.10. In the group $\widehat{\mathbb{G}_m}$, the set $\{\gamma_{\underline{m}}\}$ is a basic set of curves. What is $\sum_{\widehat{\mathbb{G}_m}}V_n[x_n]\gamma_{\underline{m}}$?

If we write the result as $\sum a_nt^n$ then $\sum a_nt^n=\sum_{\widehat{\mathbb{G}_m}}x_nt^n$. Equivalently,

$$1-\sum_{n=1}^{\infty}a_nt^n=\prod_{n=1}^{\infty}(1-x_nt^n). \tag{4.35}$$

One proves that the a_i's and the x_i's are functions of each other. We may write

$$\sum V_n[x_n]\gamma_{\underline{m}}+\sum V_n[y_n]\gamma_{\underline{m}}=\sum V_n[z_n]\gamma_{\underline{m}}, \tag{4.36}$$

and the z_n are determined by the identity

$$\prod(1-x_nt^n)\prod(1-y_nt^n)=\prod(1-z_nt^n). \tag{4.37}$$

We define:

- $W^+(K)=K^{\mathbb{N}^*}$ with the addition law $(x_n)+(y_n)=(z_n)$.
- $\widehat{W}^+(K)=K^{(\mathbb{N}^*)}$ with the addition law $(x_n)+(y_n)=(z_n)$.

REMARK 4.11. The group $W^+(K)$, canonically isomorphic to $\mathcal{C}(\widehat{\mathbb{G}_m})$, is isomorphic to the group of infinite Witt vectors over $\mathbb{Z}$. See [**48**, Section 17.1], " Lots of Witt vectors", for lots on Witt vectors. See also Section 6.1 below.

Let $\gamma_{\underline{w}}$ be the canonical curve in $\hat{W}^+$ given by $\gamma_{\underline{w}}(t)=(t,0,0,\dots)$ and define curves ϵ_i by $\epsilon_i(t)=(0,\dots,0,\underset{i}{t},0,\dots)$. In fact, $\epsilon_i=F_i\gamma_{\underline{w}}$.

THEOREM 4.12. (Representation Theorem, [**Laz**, III.4]) *There exists a canonical isomorphism*

$$\mathcal{C}(\mathcal{G})\cong\mathrm{Hom}(\widehat{W}^+,\mathcal{G}). \tag{4.38}$$

It is defined as follows: Given $\gamma \in \mathcal{C}(\mathcal{G})$ there exists a unique homomorphism of formal groups

$$u_\gamma : \widehat{W}^+ \longrightarrow \mathcal{G}, \tag{4.39}$$

such that

$$\gamma = u_\gamma \circ \gamma_{\underline{w}}. \tag{4.40}$$

COROLLARY 4.13. *There is a canonical bijection*

$$\mathrm{Cart}(K) \longrightarrow \mathcal{C}(\widehat{W}^+). \tag{4.41}$$

PROOF. Given $x \in \mathrm{Cart}(K)$ associate to it the curve $x\gamma_{\underline{w}}$ in $\mathcal{C}(\widehat{W}^+)$. For any formal group $\mathcal{G}$ and $\gamma \in \mathcal{C}(\mathcal{G})$ we have,

$$x\gamma = x(u_\gamma \circ \gamma_w) = u_\gamma \circ (x\gamma_{\underline{w}}). \tag{4.42}$$

This shows that the map

$$\mathrm{Cart}(K) \longrightarrow \mathcal{C}(\widehat{W}^+), \;\; x \mapsto x\gamma_{\underline{w}}, \tag{4.43}$$

is injective. To show it is surjective, we argue as follows: Given $\gamma' \in \mathcal{C}(\widehat{W}^+)$, define an operator x in Cart(K) by

$$x\gamma = u_\gamma \circ \gamma'. \tag{4.44}$$

We need to check that for any morphism $f : \mathcal{G} \longrightarrow \mathcal{G}'$ we have $x(f \circ \gamma) = f \circ (x\gamma)$. Indeed $x(f \circ \gamma) = u_{f\circ\gamma} \circ \gamma' = f \circ u_\gamma \circ \gamma' = f \circ (x\gamma)$. Since this holds for any morphism, we may apply that to $m : \mathcal{G} \times \mathcal{G} \longrightarrow \mathcal{G}$ to deduce that x is additive. □

We have found a bijection $\mathrm{Cart}(K) \longrightarrow \mathcal{C}(\widehat{W}^+)$ by sending x to $x\gamma_{\underline{w}}$. The curves $(\epsilon_i)_{i\in\mathbb{N}^*}$ form a basic set of curves in $\widehat{W}^+$ and hence by Proposition 4.8 every curve can be uniquely expressed as

$$\sum_{i,m\in\mathbb{N}^*} V_m[x_{m,i}]\epsilon_i = \sum_{i,m\in\mathbb{N}^*} V_m[x_{m,i}]F_i\gamma_{\underline{w}}, \;\; \forall m, (x_{m,i})_i \in K^{(\mathbb{N}^*)}. \tag{4.45}$$

COROLLARY 4.14. *The ring* Cart(K) *has the following description:*

$$\mathrm{Cart}(K) = \left\{ \sum_{i,m\in\mathbb{N}^*} V_m[x_{m,i}]F_i : \forall m \; (x_{m,i})_i \in K^{(\mathbb{N}^*)} \right\}. \tag{4.46}$$

5. Modules over Cart(K)

We have seen that for every formal group $\mathcal{G}$ the group $\mathcal{C}(\mathcal{G})$ is a module over Cart(K). These modules have certain properties which we attempt to single out now. Our motivation is to use such modules to establish an equivalence of categories between them and formal groups. The main new ingredient we need to pay attention to are topological properties.

Let $x = \sum_{i,m\in\mathbb{N}^*} V_m[x_{m,i}]F_i$ be an element of Cart(K). Put

$$\mathrm{ord}(x) = \min\{m : \exists n \; x_{m,n} \neq 0\}. \tag{5.1}$$

One may verify ([**Laz**, IV.2]) that

$$\mathrm{ord}(x \pm y) \geq \min\{\mathrm{ord}(x), \mathrm{ord}(y)\}, \; \mathrm{ord}(xy) \geq \mathrm{ord}(x), \tag{5.2}$$

$$\operatorname{ord}([a+b]-[a]-[b])>1. \tag{5.3}$$

Define ideals I_n of $\mathrm{Cart}(K)$ by

$$I_n=\{x:\operatorname{ord}(x)\geq n\}. \tag{5.4}$$

Under the topology defined by these ideals $\mathrm{Cart}(K)$ becomes a complete Hausdorff topological ring.

DEFINITION 5.1. 1. A *uniform* $\mathrm{Cart}(K)$-module is a topological continuous $\mathrm{Cart}(K)$-module C such that for any indexed set $(x_j)_{j\in J}$ in $\mathrm{Cart}(K)$ converging to zero, and for any set of elements $(\gamma_j)_{j\in J}$, we have $\sum_{j\in J}x_j\gamma_j$ converging in C.

2. For a uniform module C we let C_n be the closure of the sum of all the subgroups V_iC for $i\geq n$. Then C is a complete Hausdorff module with respect to the C_n's. Let

$$\mathbf{gr}_nC=C_n/C_{n+1}. \tag{5.5}$$

3. A *reduced* $\mathrm{Cart}(K)$-module C is a uniform module C such that the following hold:

- Its topology is the C_n topology.
- The homomorphism $\mathbf{gr}_1C\longrightarrow\mathbf{gr}_nC$ given by multiplication by V_n is bijective for every n.
- The K-module $\mathbf{gr}_1C$ is free.

The following theorem is the main theorem of this theory. It is due to P. Cartier.

THEOREM 5.2. (Main Theorem, [**Laz**, III.11]) *The category of formal groups over K with morphisms given by homomorphisms of formal groups, is equivalent to the category of reduced* $\mathrm{Cart}(K)$*-modules with morphisms given by continuous* $\mathrm{Cart}(K)$*-linear maps. The equivalence is given by*

$$\mathcal{G}\mapsto\mathcal{C}(\mathcal{G}). \tag{5.6}$$

Unfortunately, this theorem is not stated precisely enough in some early references. A fact that led to minor inaccuracies in the applications of the theory to deformations of abelian varieties.

It is of interest to understand, at least "qualitatively", how does one associates a formal group to a reduced $\mathrm{Cart}(K)$-module C. As a "simple" motivating example one may keep in mind the universal formal deformation of a supersingular elliptic curve.

Given any uniform $\mathrm{Cart}(K)$-module C, one associates to it a functor Γ from the category of formal varieties to the category of abelian groups

$$\mathcal{V}\mapsto\Gamma(\mathcal{V}). \tag{5.7}$$

The definition of Γ is motivated by the following reasoning: If C comes from a formal group $\mathcal{G}$, then $\Gamma(\mathcal{V})$ should be equivalent to $\mathrm{Mor}(\mathcal{V},\mathcal{G})$. Since the functor $\mathcal{C}$ is faithful, we have an injection

$$\mathrm{Mor}(\mathcal{V},\mathcal{G})\hookrightarrow\mathrm{Mor}(\mathcal{C}(\mathcal{V}),\mathcal{C}(\mathcal{G})). \tag{5.8}$$

Since C is to be identified with $\mathcal{C}(\mathcal{G})$, one defines $\Gamma(\mathcal{V})$ as a certain functorial subgroup of $\mathrm{Mor}(\mathcal{C}(\mathcal{V}),C)$. It consists of all the maps of the following form:

Let $(\gamma_j)_{j\in J}$ be any indexed set of elements of C and $(\pi_j)_{j\in J} : \mathcal{V} \longrightarrow D^{(J)}$ a morphism. Define a morphism $f : \mathcal{C}(\mathcal{V}) \longrightarrow C$ by

$$f(\delta) = \sum_{j\in J} \mathrm{comp}(\pi_j \circ \delta) \cdot \gamma_j, \quad \delta : D \longrightarrow \mathcal{V}. \tag{5.9}$$

One denotes f by $\sum \pi_j \star \gamma_j$.

Let us assume now further that C is reduced. A *V-basis* for C is an indexed set $(\gamma_j)_{j\in J}$ of elements of C such that the γ_j mod VC are a K-basis of the free K-module $\mathbf{gr}_1 C = C/C_2$. One then proves that every element in C may be written uniquely as $\sum_{n,j\in\mathbb{N}^*} V_{n,j}[x_{n,j}]\gamma_j$ where for every n we have $(x_{n,j})_{j\in J} \in K^{(J)}$. One shows that C is in bijective correspondence with $\mathcal{C}(D^{(J)})$ given by associating to a curve $\varphi = (\varphi_j) : D \longrightarrow D^{(J)}$ the element $\sum_{j\in J} \mathrm{comp}(\varphi_j) \cdot \gamma_j$ of C. One further shows that there is a functorial isomorphism of topological groups

$$\mathrm{Mor}(\mathcal{V}, D^{(J)}) \cong \Gamma(\mathcal{V}), \tag{5.10}$$

obtained by associating to $(\pi_j) : \mathcal{V} \longrightarrow D^{(J)}$ the element $\sum_{j\in J} \pi_j \star \gamma_j$.

6. The $\mathbb{Q}$-case

THEOREM 6.1. *(*The $\mathbb{Q}$ theorem, [**Laz**, II.3]) *Assume that K is a $\mathbb{Q}$-algebra and $\mathcal{G}, \mathcal{G}'$ are two formal groups over K. Then, to any K-linear map $u : \mathcal{T}\mathcal{G} \longrightarrow \mathcal{T}\mathcal{G}'$, there exists a unique formal group homomorphism $f : \mathcal{G} \longrightarrow \mathcal{G}'$ such that $\mathcal{T}f = u$.*

Thus, the category of formal groups over K is equivalent to the category of free K-modules. The equivalence is given by the fully-faithful functor $\mathcal{T}$. In particular, there exist unique formal group isomorphisms:

$$\log_{\mathcal{G}} : \mathcal{G} \longrightarrow (\mathcal{T}\mathcal{G})^+, \quad \exp_{\mathcal{G}} : (\mathcal{T}\mathcal{G})^+ \longrightarrow \mathcal{G}, \tag{6.1}$$

which are inverses of each other and such that

$$\mathcal{T}\log_{\mathcal{G}} = \mathrm{Id}, \quad \mathcal{T}\exp_{\mathcal{G}} = \mathrm{Id}. \tag{6.2}$$

EXAMPLE 6.2. For $\widehat{\mathbb{G}_a}$ the logarithm and exponent are of course the identity maps. For $\widehat{\mathbb{G}_m}$ we have

$$\log_{\widehat{\mathbb{G}_m}}(x) = -\log(1-x), \quad \exp_{\widehat{\mathbb{G}_m}}(x) = 1 - \exp(-x). \tag{6.3}$$

For the case of W^+ see Section 6.1. For the case of elliptic curves, see [**124**] and [**107**]. We can not resist mentioning though the following beautiful theorem connecting the arithmetic and geometry of elliptic curves.

THEOREM 6.3. (Honda, [**49**]) *Let E be an elliptic curve over $\mathbb{Q}$; let $L(E/\mathbb{Q}, s) = \sum_{n=1}^{\infty} a_n n^{-s}$ be its L-function. Put $f(x) = \sum_{n=1}^{\infty} a_n n^{-1} x^n$ and let $\mathcal{G}$ be the formal group D over $\mathbb{Q}$ with logarithm $\log_{\mathcal{G}} = f(x)$. Then G is a formal group over $\mathbb{Z}$ and is isomorphic to the formal group of the Néron model of E over $\mathbb{Z}$.*

6.1. Digression on Witt vectors. We follow [**48**, Section 17]. Define polynomials

$$w_n(x_1, \ldots, x_n) = \sum_{d|n} d x_d^{n/d} \quad \in \mathbb{Z}[x_1, \ldots, x_n]. \tag{6.4}$$

We shall write $\underline{x} = (x_1, x_2, \dots)$ and would consider w_n also as functions of $\underline{x}$. Thus, for example,

(6.5)
$$w_1(\underline{x}) = x_1,\ w_2(\underline{x}) = x_1^2 + 2x_2,\ w_3(\underline{x}) = x_1^3 + 3x_3,\ w_4(\underline{x}) = x_1^4 + 2x_2^2 + 4x_4,\ \dots.$$

One proves that there exists polynomials

$$\begin{aligned}
&\Sigma_n(x_1, \dots, x_n; y_1, \dots, y_n) \in \mathbb{Z}[x_1, \dots, x_n; y_1, \dots, y_n], \quad n = 1, 2, \dots.\\
&\Pi_n(x_1, \dots, x_n; y_1, \dots, y_n) \in \mathbb{Z}[x_1, \dots, x_n; y_1, \dots, y_n], \quad n = 1, 2, \dots.\\
&\iota_n(x_1, \dots, x_n) \in \mathbb{Z}[x_1, \dots, x_n], \quad n = 1, 2, \dots.
\end{aligned}$$

such that for every n

$$w_n(\Sigma) = w_n(\underline{x}) + w_n(\underline{y}),\ \ w_n(\Pi) = w_n(\underline{x}) \cdot w_n(\underline{y}),\ \ w_n(\iota) = -w_n, \tag{6.6}$$

where we have put

$$\Sigma = (\Sigma_1, \Sigma_2, \dots),\ \ \Pi = (\Pi_1, \Pi_2, \dots),\ \ \iota = (\iota_1, \iota_2, \dots). \tag{6.7}$$

Define a ring functor from the category of commutative rings to itself

$$W : \mathbf{Rings} \longrightarrow \mathbf{Rings}, \quad A \mapsto W(A), \tag{6.8}$$

where $W(A)$ is the Witt ring $A^{\mathbb{N}^*}$ with addition and multiplication defined by Σ, Π and i. Namely,

$$\underline{x} + \underline{y} = \Sigma(\underline{x}; \underline{y}),\ \ \underline{x} \cdot \underline{y} = \Pi(\underline{x}; \underline{y}),\ \ -\underline{x} = \iota(\underline{x}). \tag{6.9}$$

Still more visibly,

$$\begin{aligned}
&(a_1, a_2, \dots) + (b_1, b_2, \dots) = (\Sigma_1(a_1; b_1), \Sigma_2(a_1, a_2; b_1, b_2), \dots),\\
&(a_1, a_2, \dots) \cdot (b_1, b_2, \dots) = (\Pi_1(a_1; b_1), \Pi_2(a_1, a_2; b_1, b_2), \dots),\\
&-(a_1, a_2, \dots) = (\iota(a_1), \iota(a_1, a_2), \dots).
\end{aligned}$$

One may amuse himself by calculating the first of these polynomials:

$$\begin{aligned}
&\Sigma_1(x_1; y_1) = x_1 + y_1,\ \ \Sigma_2(x_1, x_2; y_1, y_2) = x_2 + y_2 - x_1 y_1, \dots.\\
&\Pi_1(x_1; y_1) = x_1 y_1,\ \ \Pi_2(x_1, x_2; y_1, y_2) = x_2 y_1^2 + y_2 x_1^2 + 2x_2 y_2, \dots.\\
&\iota_1(x_1) = -x_1,\ \ \iota_2(x_1, x_2) = -x_1^2 - x_2, \dots.
\end{aligned}$$

One can prove using Equation (6.6) that this indeed gives a ring structure. Let $\widehat{W}$ be the sub-functor, such that

$$\widehat{W}(A) = A^{(\mathbb{N}^*)}. \tag{6.10}$$

Then clearly the underlying additive group of $\widehat{W}$, denoted $\widehat{W}^+$ is a formal group. Indeed, the formal group law is non other than

$$F(\underline{x}, \underline{y}) = (\Sigma_1(\underline{x}; \underline{y}), \Sigma_2(\underline{x}; \underline{y}), \dots). \tag{6.11}$$

The map

$$\underline{x} = (x_1, x_2, \dots) \xrightarrow{\log} w(x) = (w_1(\underline{x}), w_2(\underline{x}), \dots), \tag{6.12}$$

is the logarithm of the formal group $\widehat{W}^+$.

Consider now the ring $1 + tA[[t]]$ endowed with a ring structure as follows: The usual multiplication of power series is decreed *addition*. To define *multiplication*, write formally

$$1 + x_1t + x_2t^2 + \cdots = \prod_{i=1}^{\infty}(1 - \xi_i t), \quad 1 + y_1t + y_2t^2 + \cdots = \prod_{j=1}^{\infty}(1 - \eta_j t)\ . \tag{6.13}$$

Then

$$\prod_{i,j=1}^{\infty}(1 - \xi_i\eta_j t) = 1 + P_1(x_1; y_1)t + P_2(x_1, x_2; y_1, y_2)t^2 + \dots\ . \tag{6.14}$$

Multiplication is defined by

$$(1 + x_1t + x_2t^2 + \dots) \star (1 + y_1t + y_2t^2 + \dots) = 1 + P_1(x_1; y_1)t + P_2(x_1, x_2; y_1, y_2)t^2 + \dots\ . \tag{6.15}$$

The map

$$\mathbf{E} : W(A) \longrightarrow 1 + tA[[t]], \quad \mathbf{E}(a_1, a_2, \dots) = \prod_{i=1}^{\infty}(1 - a_i t^i), \tag{6.16}$$

is an isomorphism of rings. This justifies Remark 4.11, as well as the notation.

Fix a prime p. We may then consider only "the p-part" of W and $\widehat{W}$. That is, consider only n's that are power of p. Bearing this in mind, we change notation (as is customary) and let

$$\phi_n(x_0, \cdots, x_n) = \sum_{i=0}^{n} p^i x_i^{p^{n-i}} = x_0^{p^n} + \cdots + p^{n-1}x_{n-1}^p + p^n x_n. \tag{6.17}$$

That is, $\phi_n = w_{p^n}$. Then, the polynomials $s_n = \Sigma_{p^n}$, $p_n = \Pi_{p^n}$ and $i_n = \iota_{p^n}$ define a ring structure on $W_p(A) = A^{\mathbb{N}^*}$ and $\widehat{W}_p(A) = A^{(\mathbb{N}^*)}$. This is clear from what was said above.

A fundamental result concerning $W_p(K)$ where K is a finite field of characteristic p, is that $W_p(K)$ is the unique, up to isomorphism, discrete complete valuation ring of characteristic 0 with residue field K. The maximal ideal is

$$\mathfrak{m} = \{(0, x_2, x_3, \cdots) : x_i \in K\}. \tag{6.18}$$

The map

$$K^\times \longrightarrow W_p(K), \quad x \mapsto (x, 0, 0, \dots), \tag{6.19}$$

is the *Teichmüller lift*. That is, the image of x is the unique root of unity reducing to x modulo the maximal ideal.

Since W_p is a functor, one may ask what is the functorial map $W_p(K) \longrightarrow W_p(K)$ corresponding to the Frobenius map $x \mapsto x^p$ on K? It is given just by

$${}^F(x_0, x_1, x_2, \dots) \mapsto {}^F(x_0, x_1, x_2, \dots) := (x_0^p, x_1^p, x_2^p, \dots). \tag{6.20}$$

There is another map, the *verschiebung*,

$$V : W_p(K) \longrightarrow W_p(K), \quad {}^V(x_1, x_2, \dots) = (0, x_1, x_2, \dots). \tag{6.21}$$

It is a map of the underlying additive group, but does not respect multiplication. A fundamental relation is

$$FV = VF = [p]. \tag{6.22}$$

7. Formal groups in characteristic p

If one is willing to work in characteristic p only, then the constructions above can be simplified. Let us assume henceforth that K is of positive characteristic p.

One replaces the ring $\mathrm{Cart}(K)$ by the ring $\mathrm{Cart}_p(K)$. It is the sub ring of $\mathrm{Cart}(K)$ consisting of all the sums

$$\sum_{i,n\in\mathbb{N}} V^n[x_{n,i}]F^i, \quad \forall n\ (x_{n,i})_i \in K^{(\mathbb{N})}, \tag{7.1}$$

where

$$V^n := V_{p^n}, \ \ F^i = F_{p^i}. \tag{7.2}$$

Recall the relations (let $V = V^1$ and $F = F^1$):

$$F^iF^j = F^{i+j}, \ \ V^iV^j = V^{i+j}, \ \ FV = p\cdot 1, \tag{7.3}$$

$$[a]V = V[a^p], \ \ F[a] = [a^p]F, \ \ [ab] = [a][b], \tag{7.4}$$

and the special feature of characteristic p:

$$VF = p\cdot 1. \tag{7.5}$$

Henceforth, we let $W_p(K)$ stand for the Witt vectors over K (consisting in fact of the p-power components of the previously defined $W^+(K)$). One can prove that the map

$$\mathrm{op} : W_p(K) \longrightarrow \mathrm{Cart}_p(K), \tag{7.6}$$

defined by

$$\mathrm{op}(\underline{x}) = \mathrm{op}(x_1, x_2, \dots) = \sum V^n[x_n]F^n, \tag{7.7}$$

is a ring homomorphism, and the following identities hold:

$$F^m(\mathrm{op})(\underline{x}) = \mathrm{op}({}^{F^m}\underline{x})F^m, \ \ \mathrm{op}(\underline{x})V^m = V^m\mathrm{op}({}^{F^m}\underline{x}). \tag{7.8}$$

In particular, every $\mathrm{Cart}_p(K)$*-module is canonically a* $W_p(K)$*-module via the ring homomorphism* op. We remark that the same holds for $\mathrm{Cart}(K)$ for a general ring K. The map $\mathrm{op}(x_1, x_2, \dots) = \sum V_n[x_n]F_n$ is an embedding of $W(K)$ in $\mathrm{Cart}(K)$.

One defines in the same way the notions of uniform and reduced $\mathrm{Cart}_p(K)$-modules (but see below). An element γ in a $\mathrm{Cart}_p(K)$-module is called *p-typical* if $F_n\gamma = 0$ for every n prime to p. Let

$$\mathcal{C}_p(\mathcal{G}) \subset \mathcal{C}(\mathcal{G}) \tag{7.9}$$

be the subgroup of p-typical curves.

THEOREM 7.1. (Main Local Theorem, [**Laz**, IV.7, IV.8]) *Let K be a ring of characteristic p. The functor*

$$\mathcal{G} \mapsto \mathcal{C}_p(\mathcal{G}), \tag{7.10}$$

is an equivalence of categories from the category of formal groups over K, to the category of reduced $\mathrm{Cart}_p(K)$*-modules.*

Further simplifications occurring in characteristic p are:

- A $\mathrm{Cart}_p(K)$-module C is uniform if and only if $\cap_n V^n\mathcal{C} = \{0\}$.
- A $\mathrm{Cart}_p(K)$-module C is reduced if and only if it is uniform, V is injective and C/VC is a free K-module.
- Every $\gamma \in \mathcal{C}(\mathcal{G})$ has a unique expression $\gamma = \sum_{(n,p)=1} V_n\gamma_n$ for appropriate $\gamma_n \in \mathcal{C}_p(\mathcal{G})$. This implies that $\mathcal{C}_p(\mathcal{G})/V\mathcal{C}_p(\mathcal{G}) \cong \mathcal{T}\mathcal{G}$.

REMARK 7.2. In the definition of a reduced module one requires that $\cap_i V^i C = \{0\}$. Why? The operation V is coming from the "Frobenius map",

$$R[[x_1, \ldots, x_n]] \longrightarrow R[[x_1, \ldots, x_n]], \tag{7.11}$$

taking x_i to x_i^p. This makes it clear it should be nilpotent. But note that the corresponding action on the p-divisible group is again Frobenius. Thus, in the covariant theory that we present, the operator V on the module corresponds to the Frobenius morphism on the p-divisible group. Since the p-divisible group is connected, the nilpotence of V is reflecting the fact that on a connected p-divisible group Frobenius is nilpotent.

8. Classification of p-divisible groups in characteristic p, Newton polygons and types

Recall that we are given a totally real field L, and let K stand for a characteristic p ring and k for an algebraically closed field of characteristic p. We use R to denote a general ring but whenever we say $\mathrm{Cart}_p(R)$-module, or consider $\mathcal{C}_p(\mathcal{G})$ for a formal group $\mathcal{G}$ over R, it is tacitly assumed that R has characteristic p. All formal groups in this section are finite dimensional.

DEFINITION 8.1. 1. A formal group with real multiplication by L, or simply, *formal group with RM*, is a formal group $\mathcal{G}$ over a ring R, together with an embedding of rings

$$\iota : \mathcal{O}_L \longrightarrow \mathrm{End}_R(\mathcal{G}), \tag{8.1}$$

making $\mathcal{T}\mathcal{G}$ into a locally-free $\mathcal{O}_L \otimes R$-module of rank 1.

2. A $\mathrm{Cart}(R)$-module (resp. $\mathrm{Cart}_p(R)$-module) with RM , is a reduced $\mathrm{Cart}(R)$-module (resp. $\mathrm{Cart}_p(R)$-module) C together with an embedding of rings

$$\iota : \mathcal{O}_L \longrightarrow \mathrm{End}(C) \tag{8.2}$$

(endomorphisms of topological Cartier modules), making $\mathbf{gr}_1 C$ into a locally free $\mathcal{O}_L \otimes R$-module of rank 1.

Clearly, Theorem 5.2 (resp. Theorem 7.1) implies that the functor

$$\mathcal{G} \mapsto \mathcal{C}(\mathcal{G}), \quad (\text{resp. } \mathcal{G} \mapsto \mathcal{C}_p(\mathcal{G})), \tag{8.3}$$

is an equivalence of categories between formal groups with RM and morphisms commuting with the $\mathcal{O}_L$-structure, to the category of reduced modules with morphisms of Cartier modules commuting with the $\mathcal{O}_L$-structure.

Recall that a formal group $\mathcal{G}$ over k is p-divisible, if, fixing an isomorphism $\mathcal{G} \cong D^g$, the ring D^g becomes a finite module over itself via the homomorphism $D^g \longrightarrow D^g$ induced by multiplication by p on the formal group. Equivalently, the module $\mathcal{C}_p(\mathcal{G})/p\mathcal{C}_p(\mathcal{G})$ is a finite k-module. We remark that for a p-divisible group $\mathcal{G}$ one has that $\mathcal{C}_p(\mathcal{G})$ is a free $W_p(k)$-module of finite rank.

Two p-divisible groups $\mathcal{G}_1, \mathcal{G}_2$, over k are called *isogenous* (notation: $\mathcal{G}_1 \sim \mathcal{G}_2$) if there exists a surjective homomorphism $\mathcal{G}_1 \longrightarrow \mathcal{G}_2$ with finite kernel. This is an equivalence relation. The groups $\mathcal{G}_1$ and $\mathcal{G}_2$ are isogenous if and only if $\mathcal{C}_p(\mathcal{G}_1) \otimes \mathbb{Q}$ and $\mathcal{C}_p(\mathcal{G}_2) \otimes \mathbb{Q}$ are isomorphic as $\mathrm{Cart}_p(k) \otimes \mathbb{Q}$ modules.

The classification of p-divisible formal groups over k up to isogeny was carried out by Dieudonné, and up to isomorphism by Manin in [**68**]. Note that the theory in [**68**] is contravariant, while we survey the covariant theory.

THEOREM 8.2. (Dieudonné, [**Laz**, VI.7, VI.8], [**68**, II.4]) *The category of p divisible formal groups up to isogeny, over an algebraically closed field k of characteristic p, is semi-simple. The simple objects are precisely the p-divisible groups $\mathcal{G}_{m,n}$ for $m, n \in \mathbb{N}^*$ and $(m,n) = 1$ or $(m,n) = (1,0)$.*

The group $\mathcal{G}_{m,n}$ has dimension m and is determined by its $\mathrm{Cart}_p(k)$ module $C_{m,n} = \mathcal{C}_p(\mathcal{G}_{m,n})$ where

$$C_{m,n} = \mathrm{Cart}_p(k)/(F^m - V^n). \tag{8.4}$$

REMARK 8.3. The group $\mathcal{G}_{1,0}$ is just $\widehat{\mathbb{G}_m}$ and is isomorphic to the formal group of an ordinary elliptic curve over k. The group $\mathcal{G}_{1,1}$ is isomorphic to the formal group of a supersingular elliptic curve over k. Sometimes one puts $\mathcal{G}_{0,1} := \underline{\mathbb{Q}_p/\mathbb{Z}_p}$, even though this is not a formal group in our sense.

DEFINITION 8.4. We say that $\mathcal{G}_{m,n}$ (or $\mathcal{C}_{m,n}$) has *slope* $\frac{m}{m+n}$ of *length* $m+n$. The *height* of a p-divisible group $\mathcal{G}$ is the integer h such that the p-torsion subgroup scheme of $\mathcal{G}$ has rank p^h. (Thus, the length of $C_{m,n}$ is in fact the height of $\mathcal{G}_{m,n}$). One gives $\mathcal{G}_{0,1}$ slope 1 of length 1.

If $\mathcal{G}$ is any p-divisible group over k and

$$\mathcal{G} \sim \oplus\, \mathcal{G}_{m_i,n_i} \tag{8.5}$$

(repetitions allowed), we associate to it a unique lower convex polygon in the plane, having increasing slopes, such that the polygon starts at $(0,0)$ and ends at $(\mathrm{height}(\mathcal{G}), \dim(\mathcal{G}))$, and has for every (m_i, n_i) a segment of slope $\frac{m_i}{m_i+n_i}$ of length $m_i + n_i$. We call it the *Newton polygon* of $\mathcal{G}$.

REMARK 8.5. Let $A/\mathbb{F}_q$ be a g-dimensional abelian variety (q a power of p) and let ϕ be that iterate of Frobenius satisfying $\phi(x) = x^q$. Assume that $|A[p](\overline{\mathbb{F}_q})|$ $= 1$. Then $A(p)$ is a connected p-divisible group, or, equivalently, a formal group, of dimension g and height $2g$. Its Newton polygon is symmetric in the sense that a slope λ appears with length r if and only if the slope $1-\lambda$ appears with length r. The polygon starts at $(0,0)$ and ends at $(2g, g)$.

Consider ϕ as an endomorphism of A, hence of $A(p)$, or of the free $W_p(\overline{\mathbb{F}_q})$-module $\mathcal{C}_p(A(p))$. Viewed in this way, it has a characteristic polynomial over $W_p(\overline{\mathbb{F}_q})$, whose Newton polygon *is* the Newton polygon of $A(p)$ as a formal group.

Suppose we are given a reduced $\mathrm{Cart}_p(k)$-module C. How does one find the Newton polygon?

In practice this is not easy to find! But consider a very lucky scenario: Suppose that with respect to some basis $(e_1, \dots, e_n)$ of the module C (a free $W_p(k)$-module of rank $n = \text{height}(\mathcal{G})$), the matrix $[V]$ of the operator V has entries in $W_p(\mathbb{F}_{p^m})$. The operator V^m is represented by the matrix

$$[V^m] = [V][V]^{\sigma^{-1}} \dots [V]^{\sigma^{-(m-1)}} \tag{8.6}$$

(here σ stands for Frobenius). Considering V^m as an operator on the $W_p(\mathbb{F}_{p^m})$ span of $(e_1, \dots, e_n)$, we find it is a linear operator represented by the matrix $[V^m]$. Take the Newton polygon $\mathcal{N}'$ of the characteristic polynomial of this matrix, and let $\mathcal{N}$ be the polygon obtained by dividing $\mathcal{N}'$ by m. That is, $(x, y) \in \mathcal{N}$ iff $(x, my) \in \mathcal{N}'$. Then $\mathcal{N}$ is the Newton polygon of C.

EXAMPLE 8.6. (B. Gross) Consider the field $\mathbb{F}_{p^2}$ where $p \equiv 3 \pmod 4$ and $i \in \mathbb{F}_{p^2}$ satisfies $i^2 = -1$. Suppose that V is given with respect to some basis by the matrix

$$[V] = \frac{-1}{2} \begin{pmatrix} 1-p & (p+1)i \\ (p+1)i & p-1 \end{pmatrix}. \tag{8.7}$$

The usual linear polynomial is $x^2 + p$, which would give the Newton polygon

Figure 5.

However, this is not the Newton polygon of the module. Indeed, if one changes basis by the matrix

$$N = \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}, \tag{8.8}$$

then the resulting matrix of V is

$$N[V](N^{-1})^{\sigma^{-1}} = \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix}. \tag{8.9}$$

Evidently, the Newton polygon is

Figure 6.

Note also that V^2 is given by

$$\frac{1}{2}\begin{pmatrix} p^2+1 & (p^2-1)i \\ (1-p^2)i & p^2+1 \end{pmatrix}. \tag{8.10}$$

Its linear characteristic polynomial is $(x^2 - 2(p^2+1)x + p^2)/4$ and its Newton polygon is twice the Newton polygon in *****

EXAMPLE 8.7. The module $\mathcal{C}_{2,3}$ is generated freely by $F, 1, V, V^2, V^3$. The matrix of V is

$$\begin{pmatrix} 0 & 0 & 0 & 0 & p \\ p & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}. \tag{8.11}$$

It has characteristic polynomial $x^5 - p^2$ whose Newton polygon is

Figure 7.

Let ℓ be the graph of any convex function $f : [0, h] \longrightarrow \mathbb{R}$. Say $f(h) = d$. Let $\mathcal{G}$ be a d-dimensional p-divisible formal group over R of height h. We say that the Newton polygon of $\mathcal{G}$ is *above* ℓ if for every homomorphism $\varphi : R \mapsto k$, from R to an algebraically closed field k, the Newton polygon of $\mathcal{G} \otimes_R k$ is above ℓ in the sense that no point of the polygon is strictly below ℓ (if (x, y) is a point of that polygon then $y \geq f(x)$). We remark that this property only depends on the point $\mathrm{Ker}(\varphi) \in \mathrm{Spec}(R)$.

THEOREM 8.8. (Grothendieck's Specialization Theorem, [**62**, 2.3.1]) *The set of points in* $\mathrm{Spec}(R)$ *for which the Newton polygon of* $\mathcal{G}$ *is above* ℓ *is Zariski closed.*

COROLLARY 8.9. *In case* I_p *(resp.* II_p*), for every* ℓ *there exists a closed subset* $\mathcal{N}_\ell$ *of* $\mathcal{A}(\mathbb{F}_p) = \mathcal{A} \otimes \mathbb{F}_p$ *(resp.* $\mathfrak{M}(\mu_N, \mathbb{F}) = \mathfrak{M}(\mu_N) \otimes \mathbb{F}$*) universal for the condition that the Newton polygon is above* ℓ*. That is, if* $(A, \lambda, \alpha)/R$ *(resp.* (A, ι, α)*) has the property that the Newton polygon of* $A(p)/R$ *is above* ℓ*, then the unique classifying morphism* $\mathrm{Spec}(R) \longrightarrow \mathcal{A}(\mathbb{F}_p)$ *(resp.* $\mathrm{Spec}(R) \longrightarrow \mathfrak{M}(\mu_N, \mathbb{F})$*) corresponding to* $(A, \lambda, \alpha)/R$ *(resp.* (A, ι, α)*) factors through* $\mathcal{N}_\ell$*.*

REMARK 8.10. Of course one can restrict to ℓ which is actually one of the Newton polygons appearing. Conjectures regarding those sub schemes $\mathcal{N}_\ell$ appeared for case I_p in [**92**] and for case II_p in [**39**].

Recently de Jong and Oort proved the following theorem:

THEOREM 8.11. (Purity Theorem, [**54**]) *Let R be a noetherian local ring of characteristic p and $\mathcal{G}$ a p-divisible group over R. Assume that the Newton polygon of $\mathcal{G}$ is constant over $R\backslash\{\mathfrak{m}_R\}$. Then, either* $\dim(R) \leq 1$ *or the Newton polygon of $\mathcal{G}$ is constant over R.*

Using this theorem one obtains lower bounds on $\dim(\mathcal{N}_\ell)$ and upper bounds on the dimension of components of $\mathcal{N}_\ell$ containing *supersingular* points (loc. cit, Introduction). These are points such that the Newton polygon of the associated abelian variety is a straight line from $(0,0)$ to $(2g,g)$. Among them, the easiest to describe are the superspecial points. A geometric point x of $\mathcal{A}(\mathbb{F}_p)$, or $\mathfrak{M}(\mu_N,\mathbb{F})$, is *superspecial* if the associated abelian variety has p-divisible group *isomorphic* to $\mathcal{G}_{1,1}^g$. By a theorem of Oort [**93**] an abelian variety is superspecial iff its a number is maximal. The $\mathrm{Cart}_p(k)$-module of such a p-divisible group is generated freely over $W_p(k)$ by generators $x_1, y_1, \ldots, x_g, y_g$ on which F acts by

$$\begin{pmatrix} 0 & -p & & & & & & \\ 1 & 0 & & & & & & 0 \\ & & 0 & -p & & & & \\ & & 1 & 0 & & & & \\ & & & & \ddots & & & \\ & & & & & 0 & -p & \\ & 0 & & & & 1 & 0 & \end{pmatrix}. \tag{8.12}$$

DEFINITION 8.12. Let $\mathcal{G}$ be a formal group with RM over k. We define the *type* of $\mathcal{G}$ to be the isomorphism class of the representation of $\mathcal{O}_L \otimes k$ on $\mathbb{D}_\alpha(\mathcal{G})$ where

$$\mathbb{D}_\alpha(\mathcal{G}) = \mathcal{C}_p(\mathcal{G})/(\mathcal{C}_p(\mathcal{G})V + \mathcal{C}_p(\mathcal{G})F). \tag{8.13}$$

Note that $\mathbb{D}_\alpha(\mathcal{G}) = \mathcal{T}\mathcal{G}/F\mathcal{T}\mathcal{G}$. The isomorphism class of this representation can be written as a formal sum $\sum_{(j,i)} \epsilon_{(j,i)}\sigma_{(j,i)}$, where $\epsilon_{(j,i)}$ is zero or one, or also with a vector

$$\tau(\mathcal{G}) = (\tau_1(\mathcal{G}), \ldots, \tau_r(\mathcal{G})), \tag{8.14}$$

where $\epsilon_{(j,i)} = 1$ iff $(j,i) \in \tau_j(\mathcal{G})$ (thus, $\tau_j(\mathcal{G})$ is a subset of $\{(j,i) : 1 \leq i \leq f_j\}$).

DEFINITION 8.13. Given a formal group $\mathcal{G}$ over k and given any $\tau = (\tau_1, \ldots, \tau_r)$ with τ_j a subset of $\{(j,i) : 1 \leq i \leq f_j\}$, we say that $\tau(\mathcal{G}) \geq \tau$, if for every j we have $\tau_j(\mathcal{G}) \supseteq \tau_j$.

Given a formal group $\mathcal{G}$ over an $\mathbb{F}$-algebra R, we say that $\tau(\mathcal{G}) \geq \tau$ if for every homomorphism $\varphi : R \longrightarrow k$ of R to an algebraically closed field k, we have $\tau(\mathcal{G} \times_R k) \supseteq \tau$. (This depends only on the point $\mathrm{Ker}(\varphi) \in \mathrm{Spec}(R)$).

THEOREM 8.14. (Specialization Theorem) *Given a type τ and a formal group $\mathcal{G}$ over R, the set of points of* $\mathrm{Spec}(R)$ *where $\tau(\mathcal{G}) \geq \tau$ is Zariski closed.*

PROOF. Consider the linearization $F^\sharp$ of F,

$$F^\sharp : \mathcal{T}\mathcal{G} \otimes_{R,\sigma} R \longrightarrow \mathcal{T}\mathcal{G}, \tag{8.15}$$

where $\sigma : R \longrightarrow R$ is the Frobenius and $F^\sharp(x \otimes \lambda) = \lambda F(x)$. Note that $F^\sharp(x \otimes \lambda^p) = \lambda^p F(x) = F(\lambda x) = F(\lambda x \otimes 1)$. Hence $F^\sharp$ is well defined.

We could have formulated the notion of type using $\mathcal{TG}/\mathrm{Im}(F^\sharp)$. First, for $m \in \mathcal{O}_L$ we have $F^\sharp(mx \otimes \lambda) = \lambda F(mx) = \lambda m F(x) = mF^\sharp(x \otimes \lambda)$. Thus, $F^\sharp$ is an $\mathcal{O}_L \otimes R$-linear map. If we decompose

$$\mathcal{TG} = \oplus_{i=1}^g R_i, \tag{8.16}$$

according to the decomposition

$$\mathcal{O}_L \otimes R = \oplus_{i=1}^g R_i, \tag{8.17}$$

then for every i we have

$$F^\sharp : R_{i-1} \otimes_{R,F} R \longrightarrow R_i. \tag{8.18}$$

In fact this map is given by multiplication by an element $a_i \in R$. Then the subset of $\mathrm{Spec}(R)$ such that i does not belong to the type is precisely B_a (i.e. where a_i is invertible and there "$R = a_i R$".). □

Given an abelian scheme with RM (A, λ, ι) over an $\mathbb{F}$-scheme S, we define $\tau(A) = \tau(A(p)^0)$.

Corollary 8.15. *For every τ there exists a Zariski closed subset W_τ of $\mathfrak{M}(\mu_N, \mathbb{F})$ which is universal for the property $\tau(A, \lambda, \iota) \supseteq \tau$.*

Remark 8.16. One may ask: if $\mathcal{G}$ is the formal group of an abelian variety A with RM why not consider $C/(VC + F^2C)$ etc.? The answer is that in the case of RM, the type determines over an algebraically closed field the p-torsion group scheme $A[p]$ ([**39**]). Thus no new information is gained by this generalization. In the case I_p, though, this leads to the Ekedahl-Oort stratification!

9. Mid-way summary

k – an algebraically closed field of positive characteristic p.
$\mathcal{C}_{k,p}$ – category of local artinian k-algebras with residue field k.
L – a totally real number field of degree g over $\mathbb{Q}$ in which p is inert.
$\mathbb{F}$ – a fixed field of p^g elements.
$\mathfrak{M}$ – moduli space of abelian varieties with RM by $\mathcal{O}_L$.
x – a k-rational point of $\mathfrak{M}$ parameterizing the object (A, ι).

There is an implicit μ_N-level ($N \geq 3$ structure and prime to p) whenever we deal with $\mathfrak{M}$ or x. Assume that the p-divisible group $A(p)$ is connected and let $\mathcal{G}_k$ be the corresponding divisible formal group. Then

$$\widehat{\mathcal{O}}_{\mathfrak{M}\otimes\mathbb{F},x} \cong R^{\mathrm{univ}}_{(A,\iota)} \cong R^{\mathrm{univ}}_{(A(p),\iota(p))} \cong R^{\mathrm{univ}}_{(\mathcal{G}_k,\iota_k)}. \tag{9.1}$$

Let us call the last ring simply $\mathbf{R}^{\mathbf{U}}$. In the above isomorphisms, the universal rings are with respect to deformations over objects of $\mathcal{C}_{k,p}$ (and are just the reduction mod p of the universal ring of deformations over $\mathcal{C}_k$).

It follows that there exists a $\mathrm{Cart}_p(\mathbf{R}^{\mathbf{U}})$-module $\mathbf{C}^{\mathbf{U}}$ such that for every ring R in $\mathcal{C}_{k,p}$ and a deformation $(\mathcal{G}_P, \iota_R)$ of $(\mathcal{G}_k, \iota_k)$ over R, there exists a unique ring homomorphism $\varphi : \mathbf{R}^{\mathbf{U}} \longrightarrow R$ such that $\mathbf{C}^{\mathbf{U}} \otimes_{W(\mathbf{R}^{\mathbf{U}})} W(R) \cong \mathcal{C}_p(\mathcal{G}_R)$ as $\mathcal{O}_L \otimes \mathrm{Cart}_p(R)$-modules.

Finally, we remark that since p is inert, $A(p)$ is connected if and only if A is not ordinary. This follows from observing that $A[p](k)$ is a module over the field of p^g elements $\mathcal{O}_L/(p)$.

10. Displays

Displays are a machinery developed to simplify and normalize the presentation of a reduced $\mathrm{Cart}_p(R)$-module, where R is a simple enough ring; e.g., a complete noetherian local ring. The need in finding a simplified presentation arises when one wants to study explicitly the local deformations of abelian varieties or p-divisible groups. For example, to study how the Newton polygon or type varies locally.

Let k be an algebraically closed field of characteristic $p > 0$, and let $\mathcal{G}$ be a formal group over k. There is no doubt that the Cartier-Dieudonné module of $\mathcal{G}$ is "the right thing". One of the reasons the theory works so well is the following: Let R be a perfect ring of characteristic p. Let $W_p(R)[F, V]$ be the ring consisting of all the expressions

$$\left\{ a_0 + \sum_{i=1}^{n} a_i F^i + \sum_{i=1}^{m} b_i V^i : a_i, b_i \in W_p(R) \right\}. \tag{10.1}$$

Make it into a ring by the obvious addition and multiplication determined by the relations;

$$FV = VF = p, \;\; Fa = {}^{F}aF, \;\; V\,{}^{F}a = aV. \tag{10.2}$$

Then:

- This makes $W_p(R)[F, V]$ into a ring! (perfectness is needed)
- The image of the natural map $W_p(R)[F, V] \longrightarrow \mathrm{Cart}_p(R)$ is dense.

This phenomenon is responsible for a considerable simplification in the study of $\mathrm{Cart}_p(R)$-modules.

The question is how to *extend* this notion to the case of, say, local artinian ring of characteristic p with residue field k (we do narrow our ambitions here, but that in fact suffices for the applications later). One way is, of course, to use $\mathcal{C}_p$ as described above, but it turns out that such modules are not "nice enough". A modification was found recently by Thomas Zink [**127**], following ideas of Mumford, Norman and Oort.

10.1. Basics. Let p be a prime. Let R be a characteristic p ring. Denote by

$$x \mapsto {}^{F}x, \;\; x \mapsto {}^{V}x, \tag{10.3}$$

the Frobenius and verschiebung morphisms of $W_p(R)$ respectively. Let

$$I_R = {}^{V}W_p(R) = \{(0, r_1, r_2, \dots) : r_i \in R\}. \tag{10.4}$$

Note that $R = W(R)/I_R$.

Definition 10.1. Let M and N be $W_p(R)$-modules. An additive map

$$\alpha : M \longrightarrow N \tag{10.5}$$

is called ^{F}*-linear map* if it satisfies

$$\alpha(\lambda \cdot m) = {}^{F}\lambda \cdot \alpha(m), \;\; \forall \lambda \in W_p(R), m \in M. \tag{10.6}$$

Let $\alpha^\sharp$ be the $W_p(R)$-linear map,

$$\alpha^\sharp : W_p(R) \otimes_{W_p(R),\,{}^{F}} M \longrightarrow N, \;\; \alpha^\sharp(\lambda \otimes m) = \lambda \cdot \alpha(m). \tag{10.7}$$

We say that α is an *epimorphism* (resp. *mono-morphism, isomorphism*) if $\alpha^\sharp$ is an epimorphism (resp. mono-morphism, isomorphism).

DEFINITION 10.2. A $3n$-display over R is a quadruple (P, Q, F, V^{-1}) such that:
- P is a finitely generated projective $W_p(R)$-module.
- $Q \subset P$ is a $W_p(R)$-submodule.
- $F : P \longrightarrow P$ and $V^{-1} : Q \longrightarrow P$ are ^{F}-linear maps.

The following holds:

(i) $I_R P \subset Q \subset P$ and there is a decomposition $P = L \oplus T$ as $W_p(R)$-module such that $Q = L \oplus I_R T$.

(ii) $V^{-1} : Q \longrightarrow P$ is an ^{F}-linear epimorphism.

(iii) $\forall x \in P, w \in W_p(R)$, we have

$$V^{-1}({}^V w \cdot x) = w \cdot Fx. \tag{10.8}$$

REMARK 10.3. Note that "there is no V map". Note also the identity

$$Fx = V^{-1}({}^V 1 \cdot x) \tag{10.9}$$

$(= p \cdot V^{-1}x$ if $x \in Q)$. Thus F is determined by V^{-1}.

Displays form a category. A morphism

$$\varphi : (P_1, Q_1, F_1, V_1^{-1}) \longrightarrow (P_2, Q_2, F_2, V_2^{-1}) \tag{10.10}$$

is a morphism $\varphi : P_1 \longrightarrow P_2$ of $W_p(R)$-modules such that:

$$\varphi(Q_1) \subset Q_2, \;\; \varphi F_1 = F_2 \varphi, \;\; \varphi V_1^{-1} = V_2^{-1} \varphi. \tag{10.11}$$

Assume that R is a noetherian local ring of characteristic p. Then the divisible formal groups over R correspond to reduced $\mathrm{Cart}_p(R)$-modules C such that C is a finitely generated projective $W_p(R)$-module. If R is also *perfect*, this means that C/pC is a finitely generated R-module, because in this case $I_R = pW_p(R)$. Such modules can *then* equivalently be described as free $W_p(R)$-modules M of finite rank, endowed with additive maps $F, V : M \longrightarrow M$ such that

$$F(w \cdot x) = {}^F w \cdot x, \;\; V({}^F w \cdot x) = w \cdot Vx, \;\; FV = VF = p. \tag{10.12}$$

Furthermore, M/VM and VM/pM are free R-modules and there exists an n such that $V^n M \subset pM$. The last condition is implicit for every reduced $\mathrm{Cart}_p(R)$-module M, because $\cap V^i C = \{0\}$. The equivalence rests on the fact that the natural map $W_p(R)[F, V] \longrightarrow \mathrm{Cart}_p(R)$ is injective with dense image.

We recall, mainly for the sake of completeness, the notion of a Dieudonné module. If R is a perfect ring, one denotes by $W_p(R)[F, V]$ the non-commutative polynomial ring in the variables F and V, subject to the relations

$$FV = VF = p, \;\; Fa = {}^F a F, \;\; V\,{}^F a = aV. \tag{10.13}$$

One calls a finitely generated projective $W_p(R)[F, V]$-module M which is finitely generated projective $W_p(R)$ module, a *Dieudonné module* if M/VM and VM/pM are projective R-modules.

Now let R be *any* ring of characteristic p. Let (P, Q, F, V^{-1}) be a display over R with a decomposition $P = L \oplus T$ as in Definition 10.2. After localizing on R we may

assume that P, Q, L and T are all free $W_p(R)$-modules. Choose bases $e_1, \dots, e_d$ of T and $e_{d+1}, \dots, e_h$ of L. There exists scalars $\alpha_{ij} \in W_p(R)$ such that

$$Fe_j = \sum_{i=1}^{h} \alpha_{ij} e_i, \quad j = 1, \dots, d, \tag{10.14}$$

$$V^{-1} e_j = \sum_{i=1}^{h} \alpha_{ij} e_i, \quad j = d+1, \dots, h. \tag{10.15}$$

Note that this determines $V^{-1}|_L$ and $V^{-1}|_{I_R T}$ because $V^{-1}({}^V w \cdot x) = w \cdot Fx$. Thus (10.14) determines V^{-1} and hence also F. Moreover, the matrix (α_{ij}) is invertible. This follows from

LEMMA 10.4. ([**127**, 1.4]) *The map*

$$V^{-1} \oplus F : L \oplus T \longrightarrow P, \tag{10.16}$$

given by $(x, y) \mapsto V^{-1}x + Fy$, *is an* ^{F}*-linear isomorphism.*

Conversely, given Equations (10.14) for (α_{ij}) invertible in $M_h(W_p(R))$, we can define a $3n$-display over R. Indeed, let T be the free $W_p(R)$-module on $e_1, \dots, e_d$, and let L be the free $W_p(R)$-module on $e_{d+1}, \dots, e_h$. Put

$$P = L \oplus T, \;\; Q = L \oplus I_R T, \tag{10.17}$$

and define F and V^{-1} by the additional relations:

$$Fe_j = \sum_{i=1}^{h} p\alpha_{ij} e_i, \quad j = d+1, \dots, h, \tag{10.18}$$

$$V^{-1}({}^V w e_j) = \sum_{i=1}^{h} w\alpha_{ij} e_i, \quad j = 1, \dots, d. \tag{10.19}$$

REMARK 10.5. The name "displays" is articulating the fact that in Equations (10.14) and (10.18) the maps F and V^{-1} are "displayed".

In general a V operator does not exist. The following definition attempts to define the nilpotency of V "were it to exist". We remark that a more natural (and complicated) definition can be given, without using the choice of display, which actually describes the nilpotence of some operator that always exists. See [**127**, Section 1].

DEFINITION 10.6. Let (P, Q, F, V^{-1}) be a $3n$-display and (α_{ij}) a displaying matrix as in Equation (10.14). Let

$$(\beta_{kl}) = (\alpha_{ij})^{-1}, \;\; B = (\overline{\beta_{kl}})_{d+1 \le k,l \le h}, \tag{10.20}$$

where $\overline{\beta}$ denotes $\beta \pmod{I_R}$. Let $B^{(p^i)}$ denote the matrix obtained from B by raising each coefficient to the p^i-power. We say (P, Q, F, V^{-1}) is a *display* if for some N,

$$B^{(p^N)} \dots B^{(p)} B = 0. \tag{10.21}$$

We also say then that V^{-1} satisfies the *nilpotence condition.*

REMARK 10.7. If R is perfect then V^{-1} is described by the matrix $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ with respect to the basis $pe_1, \dots, pe_d, e_{d+1}, \dots, e_h$. In this case we may also define an operator V, given with respect to the same basis by

$$((\alpha_{ij})^{-1})^{F^{-1}} = (\beta_{ij})^{F^{-1}}. \tag{10.22}$$

To know whether the operator V is nilpotent on the whole module P with respect to the p-adic topology (as is required for a Dieudonné module), we need to care only about

$$(\beta_{ij})_{d+1\le i,j\le h}^{F^{-1}} =: B^{F^{-1}} \pmod p. \tag{10.23}$$

The nilpotence is just that for some N

$$B^{F^{-1}} B^{F^{-2}} \dots B^{F^{-(N+1)}} = 0. \tag{10.24}$$

Operating by F^{N+1} we get precisely the condition above.

The proof of the following proposition is straight-forward.

PROPOSITION 10.8. *([**127**, 1.10]) The category of $3n$-displays over a perfect ring R is equivalent to the category of Dieudonné modules over R. Moreover, the displays correspond exactly to the Dieudonné modules for which V is topologically nilpotent for the p-adic topology; i.e., to those which extend to a reduced p-divisible* $\mathrm{Cart}_p(R)$*-module.*

It follows that the category of displays over a perfect ring R is equivalent to the category of divisible formal groups over R.

DEFINITION 10.9. A *$3n$-display with RM* is a $3n$-display (P, Q, F, V^{-1}) together with an embedding of rings, $\iota : \mathcal{O}_L \longrightarrow \mathrm{End}((P, Q, F, V^{-1}))$, that makes P/Q into a locally-free (on R) $\mathcal{O}_L \otimes R$-module of rank 1. A *display with RM* is a $3n$-display with RM which is a display.

Note that $I_R P \subset Q$ so the definition makes sense and in fact P/Q is isomorphic to $T/I_R T$.

10.2. Examples.

1. *The multiplicative display $\mathcal{P}_m$.* Let $\mathcal{P}_m = (P, Q, F, V^{-1})$ be defined as follows: $P = W_p(R)$, $Q = I_R$, $Fw = {}^F w$ and $V^{-1}({}^V w \cdot x) = w \cdot V^{-1}(x)$. One obtains a decomposition $P = L \oplus T$ by putting $L = \{0\}$. Thus $d = h = 1$ and the nilpotency condition holds vacuously. We remark that we have seen that $\mathcal{C}_p(\widehat{\mathbb{G}_m}) = \widehat{W}_p(R)^+$ and that justifies the name.

Let $\mathcal{P}_m^g$ be the g-fold product of $\mathcal{P}$. Write $\mathcal{P}_m^g = (P^g, Q^g, F, V^{-1})$. Say R is in $\underline{\mathbb{F} - \mathrm{Alg}}$. Then

$$P^g = W_p(R)^g \underset{\varphi}{\cong} W_p(R) \otimes_{W_p(\mathbb{F}_p)} \mathcal{O}_L. \tag{10.25}$$

This gives, via φ, an action of $\mathcal{O}_L$ on P^g, which makes it into a display with RM. Such displays are obtained from ordinary abelian varieties with RM.

2. *The superspecial display $\mathcal{P}_{sp}$.* Let $\mathcal{P}_{sp} = (P, Q, F, V^{-1})$ where:

$$P = W_p(R)e_2 \oplus W_p(R)e_1 = L \oplus T, \tag{10.26}$$

and $Q = L \oplus I_R T$. Let the matrix $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ define the display. One checks easily that the nilpotence condition holds ($B = 0$).

We define a superspecial display to be one isomorphic to $\mathcal{P}_{sp}^n$ for some n, and a superspecial display with RM is a display with RM isomorphic to $\mathcal{P}_{sp}^g$. Such displays are obtained from superspecial abelian varieties with RM.

The *standard superspecial display* is defined as follows: Let $\mathcal{P} = (P, Q, F, V^{-1})$ with displaying $2g \times 2g$ matrix

$$\begin{pmatrix} 0_2 & & & & \cdots & 0 & -1 \\ & & & & & 1 & 0 \\ 0 & -1 & 0_2 & & & & \\ 1 & 0 & & & & & \\ & & 0 & -1 & & & \\ & & 1 & 0 & & & \\ & & \vdots & & 0_2 & & \\ & & & & 0 & -1 & \\ \cdots & & \cdots & \cdots & 1 & 0 & 0_2 \end{pmatrix}. \tag{10.27}$$

(We let 0_n stand for the $n \times n$ zero matrix). The action of $\mathcal{O}_L$ is given by

$$a \mapsto \operatorname{diag}(\sigma_1(a), \sigma_1(a), \ldots, \sigma_g(a), \sigma_g(a)), \tag{10.28}$$

One can prove this is a superspecial display with RM.

We remark that if k is an algebraically closed field of characteristic p then there are finitely many isomorphism classes of superspecial abelian varieties with RM over k (in fact also in the situation I_p). One can prove that over such field every superspecial display with RM is isomorphic to the standard superspecial display. However there is usually more then one isomorphism class of superspecial abelian varieties.

The reason superspecial abelian varieties (or displays) are of such importance is the following

THEOREM 10.10. *Let $\tau \subset \{1, \ldots, g\}$. Every component of W_τ contains a superspecial point.*

For the proof see [**39**]. One may think of the superspecial points as "extreme points" of $\mathcal{M} \otimes \mathbb{F}$, which often play a role similar to the cusps.

10.3. Base change and deformations. Let $\phi : S \longrightarrow R$ be a ring homomorphism and let $\mathcal{P} = (P, Q, F, V^{-1})$ be a $3n$-display over S. Define the *base change to R* of $\mathcal{P}$, denoted $\mathcal{P}_R = (P_R, Q_R, F_R, V_R^{-1})$, as follows:

- $P_R = W_p(R) \otimes_{W_p(S)} P$.
- $Q_R = \operatorname{Ker}(W_p(R) \otimes_{W_p(S)} P \longrightarrow R \otimes_S (P/Q))$.
- $F_R = {}^F \otimes F$. That is, $F_R(\lambda \otimes x) = {}^F\lambda \otimes F(x)$.
- V_R^{-1} is defined as the unique $W_p(R)$-^{F}-linear homomorphism satisfying:

$$V_R^{-1}(w \otimes y) = {}^F w \otimes v^{-1} y, \quad w \in W_p(R), y \in Q; \tag{10.29}$$

$$V_R^{-1}({}^V w \otimes x) = w \otimes Fx, \quad x \in P. \tag{10.30}$$

We remark that if $P = L \oplus T$, then $Q_R = W_p(R) \otimes_{W_p(S)} L \oplus I_R \otimes_{W_p(S)} T$.

It is easy to verify that if $\mathcal{P}$ is a display (not merely a $3n$-display) then $\mathcal{P}_R$ is a display as well. If every element of $\mathrm{Ker}(\phi)$ is nilpotent, the converse is true.

DEFINITION 10.11. Let $\mathcal{P}_0$ be a display over R. Let $\phi : S \longrightarrow R$ be a homomorphism of rings. A *deformation* of $\mathcal{P}_0$ over S is a display $\mathcal{P}$ over S together with an isomorphism $\mathcal{P}_R \cong \mathcal{P}_0$.

10.4. The main result. We present a somewhat simplified form of Zink's theorem.

THEOREM 10.12. *([**127**, Theorem 9]) Let R be a ring of characteristic p. Assume that R is a complete local ring, or a ring such that R is an algebra of finite type over a field k.*

There exists a functor $\mathcal{BT}$ from the category of displays over R to the category of divisible formal groups over R, which is an equivalence of categories.

See [**127**, Theorem 3.2] for the definition of the functor $\mathcal{BT}$. It would suffice to know that the $\mathrm{Cart}_p(R)$ module of $\mathcal{BT}(\mathcal{P})$ is

$$\mathrm{Cart}_p(R) \otimes_{W_p(R)} \mathcal{P} / < F \otimes x - 1 \otimes Fx, V \otimes V^{-1}y - 1 \otimes y >_{x \in \mathcal{P}, y \in Q} \tag{10.31}$$

where the brackets denote the $\mathrm{Cart}_p(R)$-module generated by the specified generators.

Thus, heuristically speaking, the display $\mathcal{P}$ singles out the essential part of the Cartier-Dieudonné module.

11. The universal display

Let k be a field of characteristic p. Let $\mathcal{C}_{k,p}$ be the category of local artinian k-algebras with residue field k. We will describe the universal display over $\mathcal{C}_{k,p}$.

Let Λ be a topological ring such that the topology on Λ is given by ideals

$$\Lambda = \mathfrak{a}_0 \supset \mathfrak{a}_1 \supset \cdots \supset \mathfrak{a}_n \supset \ldots, \tag{11.1}$$

such that $\mathfrak{a}_i \mathfrak{a}_j \subset \mathfrak{a}_{i+j}$. Assume that Λ is complete and separated with respect to this topology and is of characteristic p. A $3n$-display over Λ is called a display if its base change to $\Lambda/\mathfrak{a}_i$ is a display, in the sense of Definition 10.2, for every i.

We remark that this modification is especially tailored to suit rings like the ring $k[[t_1, \ldots, t_g]]$ with $\mathfrak{a}_i$ equal to the power series in monomials of degree greater or equal to i.

DEFINITION 11.1. Let $\mathcal{P} = (P, Q, F, V^{-1})$ be a display over k. Define the *functor of deformations*

$$\mathbf{Def}_{\mathcal{P}} : \mathcal{C}_{k,p} \longrightarrow \underline{\mathrm{Sets}}, \tag{11.2}$$

sending each ring R in $\mathcal{C}_{k,p}$ to the isomorphism classes of pairs $(\tilde{\mathcal{P}}, j)$ consisting of a display $\tilde{\mathcal{P}}$ over R and an isomorphism $j : \tilde{\mathcal{P}}_k \longrightarrow \mathcal{P}$.

One can prove that this functor is pro-representable by a ring

$$\mathbf{R}^{\mathbf{U}} = k[[t_{k\ell} : 1 \leq k \leq d, d+1 \leq \ell \leq h]], \tag{11.3}$$

where $t_{k\ell}$ are free parameters and d and h are as in Equation (10.14) for $\mathcal{P}$. Let $T_{k\ell}$ be the Teichmüller lifts (or any other lifts) of $t_{k\ell}$ to $W_p(\mathbf{R}^{\mathbf{U}})$. Let $T = (T_{k\ell})$.

Let (α_{ij}) be the displaying $h \times h$ matrix for $\mathcal{P}$. Note that $(\alpha_{ij}) \in \mathrm{GL}_h(W_p(k))$ and can thus be considered in $\mathrm{GL}_h(W_p(\mathbf{R}^{\mathbf{U}}))$. Consider the matrix

$$(\alpha^{\mathbf{u}}_{ij}) = \begin{pmatrix} I_d & T \\ 0 & I_{h-d} \end{pmatrix} (\alpha_{ij}). \tag{11.4}$$

The symbol I_r stands for the $r \times r$ identity matrix. Write

$$(\alpha_{ij}) = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \tag{11.5}$$

where A is of size $d \times d$, B of size $d \times (h-d)$ etc.. Then, we may write

$$(\alpha^{\mathbf{u}}_{ij}) = \begin{pmatrix} A+TC & B+TD \\ C & D \end{pmatrix}. \tag{11.6}$$

Let $\mathcal{P}^{\mathbf{U}}$ denote the display over $\mathbf{R}^{\mathbf{U}}$ (meaning, modules over $W_p(\mathbf{R}^{\mathbf{U}})$ etc.) defined by the matrix $(\alpha^{\mathbf{U}}_{ij})$ in (11.6), then $\mathcal{P}^{\mathbf{U}}$ is the universal display (see [**127**, Section 2.2]*).*

It may be beneficial to recall what that means! It says that for every ring R in $\mathcal{C}_{k,p}$ and a deformation $(\tilde{\mathcal{P}}, j)$ over R with an isomorphism $j : \tilde{\mathcal{P}}_k \longrightarrow \mathcal{P}$, there exists a unique morphism $\phi : \mathbf{R}^{\mathbf{U}} \longrightarrow R$ such that $(\mathcal{P}^{\mathbf{U}}_{\phi,R}, j^{\mathbf{U}}_{\phi,R})$ is isomorphic to $(\tilde{\mathcal{P}}, j)$.

Remark 11.2. Using the theory of bi-extensions, Norman and Oort [**86**] proved that in situation I_p, the condition for keeping the polarization is the Riemann conditions on T; i.e., that T is symmetric. This is also explained in [**127**].

11.1. Polarization and Endomorphisms conditions. We would like to examine now the universal display in the case of RM. As usual, the totally real field L of degree g is fixed, and we assume for simplicity that p is inert in L. We consider a display $\mathcal{P} = (P, Q, F, V^{-1})$ with RM defined over an algebraically closed field k of characteristic p, such that the free $W_p(k)$ module P is a free $\mathcal{O}_L \otimes W_p(k)$ module of rank 2. These are the displays that actually come from abelian varieties with RM. We have

$$\mathcal{O}_L \otimes_{W_p(\mathbb{F}_p)} W_p(k) \cong \oplus_{i=1}^{g} W_p(k). \tag{11.7}$$

One may thus decompose P such that

$$P = \oplus_{i=1}^{g} P_i, \tag{11.8}$$

where each P_i is a free $W_p(k)$-module of rank 2, and $\mathcal{O}_L$ acts on P_i via its i-th embedding in $W_p(k)$. See also (1.7). Since the maps F and V^{-1} are ^{F}-linear, it follows that for every i

$$F(P_i) \subset P_{i+1}, \;\; V^{-1}(P_i) \subset P_{i+1}. \tag{11.9}$$

In the case at hand, both L and T of the normal decomposition are of rank g since T/I_kT is of rank g over k (where $I_k \lhd W_p(k)$ is the maximal ideal). It is important to note that the normal decomposition is not canonical in any sense. In fact, since k is perfect V exists, and we may find basis x_i, y_i to P_i such that y_i is in the image of V and the image of V (i.e. Q) is spanned by $px_1, y_1, \ldots, px_g, y_g$. Then L can be taken to be the span of the y_i's and T of the x_i's. We consider the basis

$$x_1, \ldots, x_g, y_1, \ldots, y_g. \tag{11.10}$$

The matrix (α_{ij}) giving the display of $\mathcal{P}$ with respect to *this* basis is of the form

$$(\alpha_{ij}) = \begin{pmatrix} \mathfrak{d}_2(a_2,\dots,a_g,a_1) & \mathfrak{d}_2(b_2,\dots,b_g,b_1) \\ \mathfrak{d}_2(c_2,\dots,c_g,c_1) & \mathfrak{d}_2(d_2,\dots,d_g,d_1) \end{pmatrix}, \tag{11.11}$$

where the notation $\mathfrak{d}_2$ stands for a sub-diagonal matrix. That is:

$$\mathfrak{d}_2(s_2,\dots s_g,s_1) = \begin{pmatrix} 0 & & & & s_1 \\ s_2 & 0 & & & \\ & s_3 & 0 & & \\ & & \ddots & \ddots & \\ & & & s_g & 0 \end{pmatrix}. \tag{11.12}$$

Follow now the recipe for the universal display $\mathcal{P}^{\mathbf{U}} = (P^{\mathbf{U}}, Q^{\mathbf{U}}, F^{\mathbf{U}}, V^{-1,\mathbf{U}})$ using this displaying matrix. Consider also the universal display with RM, $\mathcal{P}^{\mathbf{U},\mathbf{L}}$, which is a base change of $\mathcal{P}^{\mathbf{U}}$, obtained by dividing $\mathbf{R}^{\mathbf{U}}$ by an ideal.

To compute this ideal we argue as follows: assuming (as we do) that one knows $\mathfrak{M}\otimes\mathbb{F}$ to be non-singular of dimension g, it is enough to find *some* extension of the $\mathcal{O}_L$-action to $\mathcal{P}^{\mathbf{U}}$, and *some* ideal $\mathfrak{a}$ such that the following hold:

- Considered mod $\mathfrak{a}$, i.e. on $\mathcal{P}^{\mathbf{U}}_{\mathbf{R}^{\mathbf{U}}/\mathfrak{a}}$, this extended action of $\mathcal{O}_L$ defines a display with RM.
- The formal variety $\mathrm{Spf}(\mathbf{R}^{\mathbf{U}}/\mathfrak{a})$ is a non-singular g-dimensional sub variety of $\mathrm{Spf}(\mathbf{R}^{\mathbf{U}})$.

Thus, first write

$$\mathcal{P}^{\mathbf{U}} = \oplus_{i=1}^{g} \mathcal{P}_i^{\mathbf{U}}, \tag{11.13}$$

where each $\mathcal{P}_i^{\mathbf{U}}$ is a free $W_p(\mathbf{R}^{\mathbf{U}})$-module of rank 2, extending the decomposition in (11.8).

Secondly, choose $\mathfrak{a}$ to be the ideal $(t_{ij} : i \neq j)$. The condition that one needs to verify is that $\pmod{\mathfrak{a}}$ we have an equality:

$$\begin{pmatrix} \Sigma(a) & 0 \\ 0 & \Sigma(a) \end{pmatrix} \begin{pmatrix} A+TC & p(B+TD) \\ C & pD \end{pmatrix} = \begin{pmatrix} A+TC & p(B+TD) \\ C & pD \end{pmatrix} \begin{pmatrix} \Sigma(a) & 0 \\ 0 & \Sigma(a) \end{pmatrix}^F. \tag{11.14}$$

This is immediate to verify. Therefore, we have proven

THEOREM 11.3. *([**39**]) The universal ring for deformations of a display $\mathcal{P}$ with RM over k to displays with RM over rings of $\mathcal{C}_{k,p}$ is the ring $\mathbf{R}^{\mathbf{u},\mathbf{L}} = k[[t_1,\dots,t_g]]$. If*

$$\begin{pmatrix} \mathfrak{d}_2(a_2,\dots,a_g,a_1) & \mathfrak{d}_2(b_2,\dots,b_g,b_1) \\ \mathfrak{d}_2(c_2,\dots,c_g,c_1) & \mathfrak{d}_2(d_2,\dots,d_g,d_1) \end{pmatrix} \tag{11.15}$$

is the displaying matrix for $\mathcal{P}$, then the displaying matrix for the universal display with RM, $\mathcal{P}^{\mathbf{U},\mathbf{L}}$, can be take to be

$$\begin{pmatrix} \mathfrak{d}_2(a_2+t_2c_2,\dots,a_g+t_gc_g,a_1+t_1c_1) & \mathfrak{d}_2(b_2+t_2d_2,\dots,b_g+t_gd_g,b_1+t_1d_1) \\ \mathfrak{d}_2(c_2,\dots,c_g,c_1) & \mathfrak{d}_2(d_2,\dots,d_g,d_1) \end{pmatrix}. \tag{11.16}$$

11.2. The local structure of W_τ and Hasse invariants. We are now in position to determine the local structure of the varieties W_τ and to gain better understanding of the Hasse invariants.
Let $k \in \underline{\mathbb{F} - \mathrm{Alg}}$ and (A, ι) an abelian variety with RM over k (with μ_N-level, $N \geq 3$, that we suppress from the notation) such that $A(p)$ is connected (equivalently, since p is inert, non-ordinary). Let $\mathcal{P}$ be the corresponding display defined by a displaying matrix as in (11.15). We first notice the following

FACT 11.4. We have $i \in \tau(A)$ if and only if $a_i = 0 \pmod p$.

Now, if $a_i \neq 0 \pmod p$ (i.e. $i \notin \tau(A)$) then $a_i + t_i c_i \pmod p$ is an invertible element of $k[[t_1 \ldots, t_g]]$ and hence $a_i + T_i c_i$ is invertible in $W_p(k[[t_1, \ldots, t_g]])$. Therefore, every specialization of the universal display would have type not containing i. That is, if A has type τ every "generification" of A has type contained in τ. Moreover, it is evident that for every $\rho \subset \tau$ the variety W_ρ is defined locally at x (where x is the moduli point of (A, ι)) by the ideal $\mathfrak{a} = (t_i : i \in \rho)$. One obtains

THEOREM 11.5. *(Goren-Oort, [**39**]) The sub-varieties W_τ are non-singular varieties of pure dimension $g - |\tau|$. Furthermore, $W_\tau \cap W_\sigma = W_{\tau \cup \sigma}$.*

We now discuss the Hasse invariants. Recall first the definition (see Chapter 5, Section 3.1). Given A/R with RM, where R is an $\mathbb{F}$-algebra and a non-vanishing differential $\omega \in \mathfrak{t}^*_{A/R}$, letting $e_1, \ldots, e_g$ be the orthogonal idempotents of $\mathcal{O}_L \otimes_{\mathbb{F}_p} R$, we get a basis $(e_1\omega, \ldots, e_g\omega)$ of $\mathfrak{t}^*_{A/R}$. We have taken a dual basis (w.r.t. some $\mathcal{O}_L$-linear polarization) to $\mathfrak{t}_{A^t/R}$, say $(\eta_1, \ldots, \eta_g)$ and put

$$h_i(A, \iota, \omega) = F\eta_{i-1}/\eta_i. \tag{11.17}$$

We could have equally taken a basis $(\eta'_1, \ldots, \eta'_g)$ for $\mathfrak{t}_{A/R}$ and let $h_i(A, \iota, \omega) = F\eta'_{i-1}/\eta'_i$.

Now, $\mathfrak{t}_{A/R} \cong P/Q$ where $\mathcal{P} = (P, Q, F, V^{-1})$ is the display associated to A. The isomorphism is chosen to respect the $\mathcal{O}_L$-action and therefore (in the notation of Section 11.1) the η'_i is a multiple by an invertible element of $W_p(k)$ of x_i, at least (mod p). The same arguments may now be applied to the universal deformation of $\mathcal{P}$ given by Equation (11.16). Thus, if $\mathbf{R}^{\mathbf{U},\mathbf{L}} = k[[t_1, \ldots, t_g]]$, we get

LEMMA 11.6. *There exists $c \in W_p(\mathbf{R}^{\mathbf{U},\mathbf{L}})$ such that*

$$h_i(A, \iota, \omega) = c(a_i + T_i c_i) \pmod p. \tag{11.18}$$

COROLLARY 11.7. *We have an equality of divisors:*

$$(h_i) = W_i. \tag{11.19}$$

In particular, the divisor of h_i is reduced. Let $H = h_1 \cdots h_g$, then the divisor of H is a reduced normal crossing divisor, equal to the complement of the ordinary locus.

The following corollary may give some "intuitive feeling" to why must every component of W_τ contain a superspecial point. We recall a theorem of Bailly saying that if $\underline{\omega}$ is the sheaf of Hilbert modular forms of parallel weight 1, then there exists an r such that $\underline{\omega}^{\otimes r}$ is an ample line bundle.

COROLLARY 11.8. *(Raynaud's trick) The divisor (H) is ample. Thus every family of ordinary abelian varieties over a complete positive dimensional variety $S \longrightarrow \mathrm{Spec}(\mathbb{F})$, say $\mathcal{A} \longrightarrow S$, is isotrivial.*

PROOF. Indeed, for large enough k, we have that $rk \cdot (H)$ is the divisor of the section H^{rk} of the very ample line bundle $\underline{\omega}^{\otimes rk(p-1)}$, hence a very ample divisor. Thus, (H) is ample.

Now, the sheaf $\underline{\omega}^{\otimes r}$ extends to an ample line bundle over a compactification $\mathfrak{M}^*(\mathbb{F})$, and therefore the complement of (H) in $\mathfrak{M}^*(\mathbb{F})$, which contains the ordinary locus, is affine. Thus a complete sub-variety of $\mathfrak{M}^*(\mathbb{F})$ that is disjoint with (H) must be zero-dimensional.

Given the family $\mathcal{A} \longrightarrow S$, one may find a complete variety T and a finite morphism $\pi : T \longrightarrow S$ such that $\pi^*\mathcal{A} \longrightarrow T$ is endowed with a μ_N-level structure ($N \geq 3$). Let $\mathfrak{M}(\mu_N, \mathbb{F})$ stand for the moduli space of abelian varieties with RM and this level structure. Then, there exists a morphism $\psi : T \longrightarrow \mathfrak{M}(\mu_N, \mathbb{F})$ such that $\pi^*\mathcal{A}$ is the pull-back by ψ of the universal family over $\mathfrak{M}(\mu_N, \mathbb{F})$. But, by the above, the image of ψ is connected and zero-dimensional. That is, $\pi^*\mathcal{A}$ is constant, that is to say, $\mathcal{A} \longrightarrow S$ is isotrivial. □

APPENDIX A

Group Schemes

The problem we face in this section is to give the reader a feeling that he knows what group schemes are about, aware of the main examples, can follow the arguments concerning them that are spread all over this book, and can even make some of his own proofs for simple facts, and on the other hand to stay within a certain length limit. Since extensive treatises and several survey papers do exists (e.g. [**24**], [**35**], [**46**], [**114**], [**118**], [**94**]), we shall offer a rather peculiar choice of topics, assuming that the interested reader would consult the above references for a more complete picture. E.g., we will not prove the main theorems, but will prove all kind of exotic statements that are relevant to the issues discussed in the book.

We shall assume that the reader speaks, though not necessarily fluently, the language of schemes. This is imperative since many of the group schemes that interest us are not reduced, and hence are beyond the scope of classical algebraic geometry. Perhaps the inevitability of learning the subject is clear when one learns that the group scheme of the p-torsion of an elliptic curve of a field of characteristic p has either p geometric points (ordinary case) or non at all (supersingular case). The only way to retain the harmony of "p-torsion being of order p^2" is to consider the group scheme of p-torsion, which is indeed of order p^2, as a group scheme.

1. Some Definitions

Let $\pi : G \longrightarrow S$ be a scheme over S. One says G is a group scheme if there exist S - morphisms

$$inv : G \longrightarrow G, \quad m : G \times G \longrightarrow G, \tag{1.1}$$

and a section

$$e : S \longrightarrow G, \tag{1.2}$$

such that the following diagrams commute (compare (1.3)):

(1.3)

$$\begin{array}{ccc} G\times G\times G & \xrightarrow{m\times 1} & G\times G \\ {\scriptstyle 1\times m}\downarrow & & \downarrow{\scriptstyle m} \\ G\times G & \xrightarrow[m]{} & G \end{array} \quad , \quad \begin{array}{ccc} G & \xrightarrow{(e\circ\pi,1)} & G\times G \\ {\scriptstyle (1,e\circ\pi)}\downarrow & \searrow^{1} & \downarrow{\scriptstyle m} \\ G\times G & \xrightarrow[m]{} & G \end{array} \quad , \quad \begin{array}{ccc} G & \xrightarrow{(inv,1)} & G\times G \\ {\scriptstyle (1,inv)}\downarrow & \searrow^{e\circ\pi} & \downarrow{\scriptstyle m} \\ G\times G & \xrightarrow[m]{} & G \end{array} .$$

As schemes are completely determined by their functor of points, an equivalent definition is: The functor of point h_G of G gives a functor

$$h_G : \mathbf{Sch}_S \longrightarrow \underline{\text{Groups}}. \tag{1.4}$$

It is often this last definition which is easily verified.

A group scheme $\pi : G \longrightarrow S$ is called *affine* if the morphism π is affine. It means that for every open affine $U = \mathrm{Spec}(R)$ of S we have $G|_U := G \times_S U$ is affine, say equal to $\mathrm{Spec}(R[x_1, \ldots, x_{r(U)}]/I_U)$. In that case the group law is determined by morphisms that are *R-algebra homomorphisms* (compare Chapter 1 (1.4)):

$$\widetilde{m} : R[x_1, \ldots, x_{r(U)}]/I_U \longrightarrow R[x_1, \ldots, x_{r(U)}]/I_U \otimes_R R[x_1, \ldots, x_{r(U)}]/I_U, \tag{1.5}$$

$$\widetilde{inv} : R[x_1, \ldots, x_{r(U)}]/I_U \longrightarrow R[x_1, \ldots, x_{r(U)}]/I_U, \tag{1.6}$$

$$\widetilde{e} : R[x_1, \ldots, x_{r(U)}]/I_U \longrightarrow R. \tag{1.7}$$

The kernel of $\widetilde{e}$ is called the *augmentation ideal*.

A group scheme $\pi : G \longrightarrow S$ is called *flat* if the morphism π is flat. That means that one can cover S by a open affine sets $U = \mathrm{Spec}(R)$ such that over each U there exists a covering of G by open affine sets $V = \mathrm{Spec}(R')$ such that R' is a flat R-algebra. Recall that this means that the functor $M \mapsto M \otimes_R R'$, from R-modules to R'-modules is exact. Though probably not clear from the definition, this means that G varies continuously over S. See [**78**, Chapter III.10].

A group scheme $\pi : G \longrightarrow S$ is called *finite* if π is a finite morphism. Thus, locally on S, one can write G as $\mathrm{Spec}(R[x_1, \ldots, x_{r(U)}]) \longrightarrow \mathrm{Spec}(R)$ and the module $R[x_1, \ldots, x_{r(U)}]$ is a finite R-module. If G is also flat, then this R module is locally free of a certain rank. This rank is constant if S is connected. In general, given a finite flat group scheme G over S, we assume that the rank is constant and call it the *rank of* G.

A *subgroup scheme* H of a group scheme $\pi : G \longrightarrow S$ is a closed subscheme that is a group scheme under the morphisms induced from those of G. Suppose that $\pi : G \longrightarrow S$ is an affine group scheme and that $S = \mathrm{Spec}(R)$ is affine. Write $G = \mathrm{Spec}(R_G)$ for a suitable R algebra R_G, and let I_G be the augmentation ideal of G. Let H be a subgroup scheme, then $H = \mathrm{Spec}(R_G/J)$ for a some ideal J. The properties forced on J are that $J \subset I_G$, that $\widetilde{m}(J) \subset J \otimes R_G + R_G \otimes J$ and that $\widetilde{inv}(J) \subset J$. Conversely, every ideal J with such properties defines a closed subgroup scheme.

Let $G' = \mathrm{Spec}(R_{G'})$ be another group scheme affine over S. Giving a homomorphism $f : G \longrightarrow G'$ of S-group schemes is equivalent to giving a homomorphism $\phi : R_{G'} \longrightarrow R_G$ that commutes with the maps $\widetilde{m}, \widetilde{inv}$ and $\widetilde{e}$. The *kernel* of f is defined as the fibre product

$$\begin{array}{ccc} \mathrm{Ker}(f) & \longrightarrow & G \\ \downarrow & & \downarrow f \\ S & \xrightarrow{e} & G' \end{array} \quad . \tag{1.8}$$

It is always a subgroup scheme. Even not in the affine case. In the affine case, the ideal J of R_G defining $\mathrm{Ker}(f)$ is $\phi(I_{G'})R_G$.

The notion of a quotient group scheme is much harder. Even if one attempts to divide by a normal subgroup scheme. We remark that the quotient by a finite flat group scheme always exists. ****

2. Digression on Frobenius and Verschiebung

Let S be a scheme of characteristic p. That is, $p = 0$ in the structure sheaf $\mathcal{O}_S$ of S. There is then a morphism $\mathrm{Fr}^{abs} : S \longrightarrow S$, the *absolute Frobenius*, that is given as the identity map on the underlying topological space of S and as the map of raising to the p-power on the structure sheaf $\mathcal{O}_S$.

Let $f : X \longrightarrow S$ be a scheme over S. Define $X^{(p)}$ as a fibre product in the following cartesian diagram

$$\begin{array}{ccc} X^{(p)} & \longrightarrow & X \\ \downarrow & & \downarrow f \\ S & \xrightarrow{\mathrm{Fr}^{abs}} & S \end{array} . \tag{2.1}$$

The operation $X \mapsto X^{(p)}$ is a covariant *functor* from the category of S schemes to itself.

Exercise 2.1. Determine how to define the Frobenius of a morphism and verify that we indeed get a functor. Is it an exact functor? faithful?

The commutative diagram

$$\begin{array}{ccc} X & \xrightarrow{\mathrm{Fr}^{abs}} & X \\ \downarrow f & & \downarrow f \\ S & \xrightarrow{\mathrm{Fr}^{abs}} & S \end{array} . \tag{2.2}$$

produces by the universal property of fibre product a morphism $\mathrm{Fr} = \mathrm{Fr}_X$, called the Frobenius morphism,

$$\mathrm{Fr} : X \longrightarrow X^{(p)}. \tag{2.3}$$

Note that it is a *morphism of S schemes*.

For example, let k be a perfect field of characteristic p, $S = \mathrm{Spec}(k)$, and let X be the scheme $\mathrm{Spec}(k[x_1, \ldots, x_n]/(f_1, \ldots, f_m))$. Then the scheme $X^{(p)}$ is given by $\mathrm{Spec}(k[x_1, \ldots, x_n]/(g_1, \ldots, g_m))$ where g_i is obtained from f_i by raising each coefficient of f_i to the p power. (In particular, if f_i are all in $\mathbb{F}_p$ then $X^{(p)} \cong X$). The morphism $\mathrm{Fr} : X \longrightarrow X^{(p)}$ is given by the homomorphism of k-algebra

$$k[x_1, \ldots, x_n]/(g_1, \ldots, g_m) \longrightarrow k[x_1, \ldots, x_n]/(f_1, \ldots, f_m), \tag{2.4}$$

determined uniquely by $x_i \mapsto x_i^p$ for $i = 1, \ldots, n$. In terms of the functor of points, the morphism $\mathrm{Fr} : X \longrightarrow X^{(p)}$ corresponds to

$$h_X \mapsto h_{X^{(p)}} \tag{2.5}$$

given for any k-algebra R by

$$\begin{aligned} h_X(R) = \{(r_1, \ldots, r_n) : f_i(r_1, \ldots, r_n) = 0 \ \forall i\} \\ \mapsto h_{X^{(p)}}(R) = \{(r_1, \ldots, r_n) : g_i(r_1, \ldots, r_n) = 0 \ \forall i\} \end{aligned} \tag{2.6}$$

by

$$(r_1, \ldots, r_n) \mapsto (r_1^p, \ldots, r_n^p). \tag{2.7}$$

The functoriality of the construction easily gives that if $X \longrightarrow S$ is a group scheme then so is $X^{(p)} \longrightarrow S$ and $\mathrm{Fr} : X \longrightarrow X^{(p)}$ is a group homomorphism.
Returning to the general case, assume that G is a finite flat commutative group scheme or an abelian scheme (see below) over S. Then duality theory provides one with a morphism $\mathrm{Ver} : G^{(p)} \longrightarrow G$ called the Verschiebung morphism. One considers the morphism $\mathrm{Fr} : G^t \longrightarrow (G^t)^{(p)}$. Here G^t is the dual group scheme (see below) if G is a finite flat group scheme, or the dual abelian scheme. Upon dualizing we obtain a morphism $\mathrm{Ver} = \mathrm{Ver}_G$

$$\mathrm{Ver} := \mathrm{Fr}^t : G^{(p)} \longrightarrow G. \tag{2.8}$$

The main property of Ver is that it is a group homomorphism satisfying

$$\mathrm{Fr}_G \circ \mathrm{Ver}_{G^{(p)}} = [p_{G^{(p)}}], \;\; \mathrm{Ver}_{G^{(p)}} \circ \mathrm{Fr}_G = [p_G]. \tag{2.9}$$

3. Important Examples

1. The multiplicative group $\mathbb{G}_m$. For every ring R

$$\mathbb{G}_{m/R} := \mathrm{Spec}(R[x, x^{-1}]), \tag{3.1}$$

with

$$\widetilde{m}(x) = x \otimes x, \;\; \widetilde{inv}(x) = x^{-1}, \;\; \widetilde{e}(x) = 1. \tag{3.2}$$

It is the group scheme associating to any R-algebra T the group $T^\times$ of invertible elements in T. The augmentation ideal is generated by $x - 1$. For every R-algebra R' we have $\mathbb{G}_{m/R'} = \mathbb{G}_{m/R} \times_{\mathrm{Spec}(R)} \mathrm{Spec}(R')$ – a feature of many of the examples below.
2. The roots of unity μ_N. Let N be a positive integer. For every ring R we define the group of N-th roots of unity by

$$\mu_N = \mathrm{Spec}(R[x]/(x^N - 1)), \tag{3.3}$$

with

$$\widetilde{m}(x) = x \otimes x, \;\; \widetilde{inv}(x) = x^{-1}, \;\; \widetilde{e}(x) = 1. \tag{3.4}$$

It is the group scheme associating to any R-algebra R' the multiplicative group $\{\zeta \in R' : \zeta^N = 1\}$ (the N-th roots of unity in R'. Or rather, the roots of unity of order N, if we want to point out that there might be more, or less, than N of them). The group μ_N is the kernel of the homomorphism $\mathbb{G}_m \longrightarrow \mathbb{G}_m$ given by $x \mapsto x^N$, or, in terms of points, for every R-algebra R'

$$\mu_N(R') \longrightarrow \mu_N(R'), \;\; \zeta \mapsto \zeta^N. \tag{3.5}$$

Note that by our recipe for kernels μ_N is to be defined by the image of $I_{\mathbb{G}_m}$, which is indeed the case. We further remark that if R has characteristic p then μ_p is the kernel of $\mathrm{Fr} : \mathbb{G}_{m/R} \longrightarrow \mathbb{G}_{m/R}$. More generally, $\mu_{p^n} = \mathrm{Ker}(\mathrm{Fr}^n)$.

Exercise 3.1. Let N be prime. Prove that $\mu_{N/R}$ has no non-proper subgroup schemes. Note that for many R' the abstract group $\mu_N(R')$ is not simple.

3. The additive group $\mathbb{G}_a$. For every ring R define the group scheme $\mathbb{G}_a$ as

$$\mathbb{G}_a := \mathrm{Spec}(R[x]), \tag{3.6}$$

with

$$\widetilde{m}(x) = x \otimes 1 + 1 \otimes x, \;\; \widetilde{inv}(x) = -x, \;\; \widetilde{e}(x) = 0. \tag{3.7}$$

It is the group scheme assigning to every R-algebra R' the underlying additive group of R'. In general the homomorphisms of $\mathbb{G}_a$ are just the maps induced by $x \mapsto rx$ for $r \in R$ that include the multiplication by n maps $x \mapsto nx$. Assume however that R has characteristic p a prime number. Then the map of raising-to-the-p-power, given on the coordinate ring by $x \mapsto x^p$, is a homomorphism of groups. It is in fact the Frobenius morphism defined above.

Thus, if R if a perfect ring of characteristic p, $\mathrm{End}(\mathbb{G}_a/R) \supset R\{\tau\}$ – the non commutative ring of polynomials in the variable τ. Every element f of $R\{\tau\}$ has a unique expression of the form $f = r_0 + r_1\tau + \cdots + r_n\tau^n$ and $\tau r = r^p\tau$. The action of f on $\mathrm{End}(\mathbb{G}_a/R)$ is given by

$$x \mapsto r_0 x + r_1 x^p + \cdots + r_n x^{p^n}. \tag{3.8}$$

This structure is fundamental to the theory of Drinfeld modules. See [**28**].

4. The group α_{p^r}. Let R be a ring of characteristic p. We define the group scheme $\alpha_{p^r/R}$ as the kernel of $\mathrm{Fr}^r : \mathbb{G}_a \longrightarrow \mathbb{G}_a$. Thus, by our recipe for kernels,

$$\alpha_{p^r/R} = \mathrm{Spec}(R[x]/(x^{p^r}), \tag{3.9}$$

with

$$\widetilde{m}(x) = x \otimes 1 + 1 \otimes x, \;\; \widetilde{inv}(x) = -x, \;\; \widetilde{e}(x) = 0. \tag{3.10}$$

It associates to every R-algebra R' the additive group of nilpotent elements of order p^r of R'. That is, $\{a \in R' : a^{p^r} = 0\}$.

Exercise 3.2. Let R be a field of characteristic p. Prove that $\mathrm{End}(\alpha_{p/R}) = R$.

5. The group GL_n. For notational simplicity we just define GL_2. For every R, we let

$$\mathrm{GL}_{2/R} := \mathrm{Spec}(R[a, b, c, d, (ad - bc)^{-1}]), \tag{3.11}$$

with $\widetilde{m}$ given by

$$\widetilde{m}(a) = a \otimes a + b \otimes c, \;\; \widetilde{m}(b) = a \otimes b + b \otimes d, \tag{3.12}$$

$$\widetilde{m}(c) = c \otimes a + d \otimes c, \;\; \widetilde{m}(d) = c \otimes b + d \otimes d, \tag{3.13}$$

with $\widetilde{inv}$ given by

$$\widetilde{inv}(a) = d(ad - bc)^{-1}, \;\; \widetilde{inv}(b) = -b(ad - bc)^{-1}, \tag{3.14}$$

$$\widetilde{inv}(c) = -c(ad - bc)^{-1}, \;\; \widetilde{inv}(d) = a(ad - bc)^{-1}, \tag{3.15}$$

and with $\widetilde{e}$ given by

$$\widetilde{e}(a) = 1, \;\; \widetilde{e}(b) = 0, \;\; \widetilde{e}(c) = 0, \;\; \widetilde{e}(d) = 1. \tag{3.16}$$

It is the group scheme associating to each R-algebra R' the group of 2×2 invertible matrices with entries in R.

The reader is well acquainted with the group homomorphism

$$\det : \mathrm{GL}_2 \longrightarrow \mathbb{G}_m \tag{3.17}$$

given by the determinant.

Exercise 3.3. Write this homomorphism in terms of the coordinate rings. The kernel is a the group scheme SL_2. What is the ideal defining it?

6. A non commutative group scheme of order p^2. Let R be a ring of characteristic p. We define a subgroup scheme of GL_2 in functorial terms. For every R-algebra R' it is given by the matrices

$$\begin{pmatrix} \zeta & \alpha \\ 0 & 1 \end{pmatrix}, \ \ \zeta \in \mu_p(R'), \alpha \in \alpha_p(R'). \tag{3.18}$$

EXERCISE 3.4. Check that this a group scheme and write it as an affine group scheme. Prove it is of rank p^2 and non commutative.

EXERCISE$^\bigstar$ 3.5. Prove that this group scheme is isomorphic to $\mu_p \ltimes \alpha_p$. For that you have to first make sense of the last expression.

7. The constant group scheme $\underline{\Gamma}$. Let Γ be a finite abelian group in the usual sense of freshmen algebra course. Let R be a ring and $S = \mathrm{Spec}(R)$. We define the constant group ring $\underline{\Gamma}$ defined by Γ as

$$\underline{\Gamma} = \coprod_{\gamma \in \Gamma} (\mathrm{Spec}(R))_\gamma = \mathrm{Spec} \underset{\gamma \in \Gamma}{\oplus} R_\gamma = \mathrm{Spec}\ R^\Gamma. \tag{3.19}$$

This defines an S-scheme $\pi : \underline{\Gamma} \longrightarrow S$. Suppose that S is connected and T is a connected S-scheme. Then $\underline{\Gamma}(T) = \Gamma$. This explains the name "constant".

We may identify R^Γ with $R[\Gamma] := \{\sum a_\gamma \gamma : a_\gamma \in R\}$. This identification sends $\gamma \in R[\Gamma]$ to the delta function at γ – an element of R^Γ. Therefore, the multiplication law induced on $R[\Gamma]$ is *not* the usual one of the group ring, but rather

$$(\sum a_\gamma \gamma)(\sum b_\gamma \gamma) = \sum a_\gamma b_\gamma \gamma. \tag{3.20}$$

To emphasize that we shall write $R[\underline{\Gamma}]$. So far what we said holds for every finite set. The group scheme structure on $\underline{\Gamma}$ comes from the group structure on Γ. One finds that the comorphisms are

$$\widetilde{m}(\gamma) = \sum_{\delta \in \Gamma} \gamma\delta \otimes \delta^{-1}, \ \ \widetilde{inv}(\gamma) = \gamma^{-1}, \ \ \widetilde{e}(\gamma) = 0. \tag{3.21}$$

EXERCISE 3.6. Given a finite flat commutative group scheme $\pi : G \longrightarrow S$, one can define a *dual* group scheme. The construction being local on the base, we may restrict to $S = \mathrm{Spec}(R)$. Then G is given by an R-algebra, say T. It comes equipped with co-multiplication, co-inverse and augmentation maps $\widetilde{m} : T \longrightarrow T \otimes_R T$, $\widetilde{inv} : T \longrightarrow T$ and $\widetilde{e} : T \longrightarrow R$. Let us denote the multiplication by $\mu : T \otimes T \longrightarrow T$ and the structure map by $\epsilon : R \longrightarrow T$.

Consider now the finite R-module $T^* := \mathrm{Hom}_R(T, R)$. Show that the map $\widetilde{m}$ induces multiplication on T^* and that T^* becomes an R-algebra with structure map induced from $\widetilde{e}$. Show that there is a natural group structure on $\mathrm{Spec}(T^*)$ for which co-multiplication is induced by μ, co-inverse is induced by $\widetilde{inv}$ and augmentation (co-unit) is induced by ϵ. The group scheme $\mathrm{Spec}(T^*)$ is called the dual group scheme to G and is usually denoted $G^*, G^t, \hat{G}$ or $G^\vee$.

The adjective "dual" is justified in that that there is a canonical perfect pairing

$$G \times G^t \longrightarrow \mathbb{G}_m, \tag{3.22}$$

and $(G^t)^t$ is naturally isomorphic to G. See ****.

EXERCISE 3.7. Show that the dual group scheme of $\underline{\Gamma}$ is the diagonalizable group scheme, discussed extensively in Chapter 1, Section 1 (but we allow p-torsion now), whose coordinate ring is $R[\Gamma]$. Note that this is the same set as the coordinate ring $R[\underline{\Gamma}]$ but now multiplication is the "usual" multiplication in a group ring:

$$(\sum_{\gamma\in\Gamma} a_\gamma\gamma)(\sum_{\gamma\in\Gamma} b_\gamma\gamma) = \sum_{\gamma\in\Gamma}\sum_{\delta\in\Gamma} a_{\gamma\delta}b_{\delta^{-1}}\gamma, \tag{3.23}$$

while co-multiplication is given by $\gamma \mapsto \gamma\otimes\gamma$.

EXERCISE 3.8. Show that μ_N and $\underline{\mathbb{Z}/n\mathbb{Z}}$ are dual to one another.

8. Étale group schemes. Recall that a morphism of schemes $\pi : T \longrightarrow S$ is called *étale* if it is finite, flat and unramified. It means that locally it is of the form $\mathrm{Spec}(R[x]/(f(x))) \longrightarrow \mathrm{Spec}(R)$ for a separable polynomial f. Heuristically, this is a concept that puts together the notion of a topological covering map and a separable field extension. See [**71**] for more on étale morphisms.

A group scheme $\pi : G \longrightarrow S$ is called étale if the morphism π is étale. The main fact one employs about étale group schemes is that after a suitable base change they become constant group schemes. The group μ_N is an étale group scheme if and only if the characteristic of every geometric point of S is prime to N.

EXERCISE 3.9. Let R be a field of characteristic p and let ℓ be a prime. Prove that the group scheme $\mu_{\ell/R}$ is étale over R if and only if $p \neq \ell$.

Suppose that R is a field, $S = \mathrm{Spec}(R)$ and K is an algebraic closure of R. Then the category of étale group schemes is equivalent to the category of finite $Gal(K/R)$ sets. See [**114**] or [**71**].

9. The p-torsion group scheme $A[p]$. Let $\pi : A \longrightarrow S$ be an abelian scheme. That is $\pi : A \longrightarrow S$ is a group scheme, the morphism π is proper, flat, with geometrically connected fibres. An abelian scheme is always commutative. It should be thought of as a continuously varying family of abelian varieties (possibly over fields of different characteristics). A typical example is the relative Jacobian. If $C \longrightarrow S$ is a family of curves, then one can put their Jacobian varieties together to one abelian scheme $\pi : A \longrightarrow S$ whose fibres are the Jacobian varieties of the corresponding curves.

Let $\pi : A \longrightarrow S$ be an abelian scheme of relative dimension g. For every integer n we denote by $[n]$ the *multiplication by n map*. It is a proper flat morphism (see [**83**]) and its kernel is a finite flat group scheme of order n^{2g} that is denoted $A[n]$. Let S_0 be the open subscheme of S where the primes dividing n are invertible. Then $A[n]$ is étale over S_0, and S_0 is maximal with such property. That is, if k is a field of characteristic p and A/k is an abelian variety then $A[p]$ is never étale. In fact, its largest étale quotient is of order $\leq p^g$. If equality exists, one says that A is *ordinary*.

We provide some examples of the structure of p-torsion of a g-dimensional abelian variety A over an algebraically closed field of characteristic p.

EXAMPLE 3.10. $g = 1$. Every elliptic curve is automatically principally polarized. This implies that $A[p]$ is self-dual. There are two possibilities:

- A is ordinary elliptic curve:

$$A[p] \cong \mu_p \oplus \underline{\mathbb{Z}/p\mathbb{Z}}. \tag{3.24}$$

Here μ_p is the kernel of Frobenius and $\underline{\mathbb{Z}/p\mathbb{Z}}$ is the kernel of Verschiebung. We remark here that for any abelian variety of dimension g in characteristic p Frobenius has degree p^g.

- A is not an ordinary elliptic curve. Then A is called a supersingular elliptic curve. One has a non-split exact sequence

$$0 \longrightarrow \alpha_p \longrightarrow A[p] \longrightarrow \alpha_p \longrightarrow 0. \tag{3.25}$$

The embedded α_p is unique and is both the kernel of Frobenius and Verschiebung. One has in fact $\mathrm{Fr} = -\mathrm{Ver}$. For every two supersingular elliptic curves A_1, A_2 over k (algebraically closed !) we have $A_1[p] \cong A_2[p]$. Thus we shall denote this group scheme by M.

EXAMPLE 3.11. $g = 2$. We assume that A is principally polarized, hence $A[p]$ is self-dual. There are four possibilities:

- A is ordinary:

$$A[p] \cong \mu_p^2 \oplus \underline{\mathbb{Z}/p\mathbb{Z}}^2. \tag{3.26}$$

Here μ_p^2 is the kernel of Frobenius and $\underline{\mathbb{Z}/p\mathbb{Z}}^2$ is the kernel of Verschiebung.

- A has étale part of order p. In this case, as in fact forced by self-duality, we have

$$A[p] \cong M \oplus \mu_p \oplus \underline{\mathbb{Z}/p\mathbb{Z}}, \tag{3.27}$$

where M is the p-torsion of a supersingular elliptic curve. Thus $A[p]$ contains a unique α_p. In this case, the kernel of Frobenius is $\alpha_p \oplus \mu_p$ and the kernel of Verschiebung is $\alpha_p \oplus \underline{\mathbb{Z}/p\mathbb{Z}}$.

- $A[p]$ has no étale part. In this case A is supersingular (but be careful: for $g \geq 3$ having no physical p-torsion (i.e., trivial étale quotient) does not imply super-singularity, though super-singularity implies no physical p-torsion). There are two possibilities:

(i) A is superspecial. That is, A is isomorphic to a product of supersingular elliptic curves. In this case

$$A[p] \cong M^2. \tag{3.28}$$

Note that $\alpha_p \oplus \alpha_p$ embeds in $A[p]$ and is in fact the kernel of both maps Frobenius and Verschiebung.

(ii) A is not superspecial. We remark that A is always isogenous to a product of two supersingular elliptic curves. In this case one has a filtration

$$H \subset G \subset A[p], \tag{3.29}$$

where $H \cong \alpha_p$, where $G/H \cong \alpha_p \oplus \alpha_p$, and where $A[p]/G \cong \alpha_p$. The kernel of Frobenius G_1 and the kernel of Verschiebung G_2 are contained in G and we have an exact sequence

$$0 \longrightarrow H \longrightarrow G_1 \oplus G_2 \longrightarrow G \longrightarrow 0. \tag{3.30}$$

We remark that neither G_1 nor G_2 are isomorphic to M. The group scheme G_1 is killed by Fr and Ver^2, the group scheme G_2 is killed by Fr^2 and Ver and G_1 is dual to G_2.

4. The Basic Exact Sequence

Let $(R, \mathfrak{m})$ be a henselian local ring. Recall that this means that Hensel's lemma holds in R. That is, if $f(x) \in R[x]$ and $\alpha_0 \in R$ are such that $f(\alpha_0) \equiv 0 \pmod{\mathfrak{m}}$ and $f'(\alpha_0) \not\equiv 0 \pmod{\mathfrak{m}}$ then there exists an $\alpha \in R$ such that $f(\alpha) = 0$ and $\alpha \equiv \alpha_0 \pmod{\mathfrak{m}}$. Examples include fields and complete local rings.

Let $S = \mathrm{Spec}(R)$ and let $\pi : G \longrightarrow S$ be a finite flat group scheme. Then there exists a canonical exact sequence

$$0 \longrightarrow G^0 \longrightarrow G \longrightarrow G^{\text{ét}} \longrightarrow 0. \tag{4.1}$$

In this sequence G^0 is the connected component of the identity. The group $G^{\text{ét}}$ is étale and is in fact the largest étale quotient of G. If R is a perfect field then the sequence splits and G is a semi-direct product $G^0 \rtimes G^{\text{ét}}$. For proofs see [**114**]. If R is a perfect field and G is commutative then we may decompose further G^0 and $G^{\text{ét}}$. The procedure being similar, we explicate only the case of G^0. One consider the dual group scheme $H = (G^0)^t$. It can be decomposed as $H = H^0 \times H^{\text{ét}}$. Dualizing, we get $G^0 = G^{0-0} \times G^{0-\text{ét}}$ where G^{0-0} is connected with connected dual, $G^{0-\text{ét}}$ is connected with étale dual. Similarly, $G^{\text{ét}}$ decomposes into a direct sum of an étale group with connected dual $G^{\text{ét-0}}$ and an étale group with étale dual $G^{\text{ét-ét}}$. All together

$$G \cong G^{0-0} \times G^{0-\text{ét}} \times G^{\text{ét-0}} \times G^{\text{ét-ét}}. \tag{4.2}$$

Using this decomposition, the category of finite flat commutative group schemes over a perfect field R decomposes into a direct sum of four categories: connected groups with connected dual, connected groups with étale dual, étale groups with connected dual, and étale groups with étale dual.

We provide some further remarks about connected group schemes. Let R be a perfect field and G a finite commutative connected group scheme over R. Then the underlying topological space of G consists of only one point, equivalently G is the spectrum of a local ring. Still equivalently, G has a unique geometric point. The last property also proves our claim. If G has more then one geometric point then its étale quotient is not trivial.

If R is a field of characteristic zero one can show that every connected group scheme is the trivial group. Assume now that R is a field of characteristic p. Then one can "effectively" construct the connected component of the identity in G. Let $G = \mathrm{Spec}(T)$ for an R-algebra T, and let I_G be the augmentation ideal. Let $I_G^{(p^b)}$ be the ideal generated by all p^b powers of elements of I (usually strictly included in I^{p^b}). Let $I_G^\infty = \cap_b I_G^{(p^b)}$. Then

$$G^0 = \mathrm{Spec}(T/I_G^\infty). \tag{4.3}$$

Equivalently, let $G[\mathrm{Fr}^b]$ denote the kernel of the Frobenius morphism iterated b times, $\mathrm{Fr} : G \longrightarrow G^{(p^b)}$. Then

$$G^0 = \bigcup_b G[\mathrm{Fr}^b]. \tag{4.4}$$

5. Group Schemes over a Perfect Field of Characteristic p

In this section we make the following standing assumptions: k is a perfect field of characteristic p; $\pi : G \longrightarrow \mathrm{Spec}(k)$ is a finite commutative group scheme of order p^g for some g.

The main tool in studying a group scheme like G is its Dieudonné module. Here we give the recipe of covariant Dieudonné modules. It differs from the theory exposed in [**24**] by taking duals.

Consider the non-commutative ring $A = W(k)[F, V]$, where $W(k)$ is the ring of infinite Witt vectors over k with the Frobenius automorphism σ, F and V are variables and

$$FV = VF = p, \quad F\lambda = \lambda^\sigma F, \quad \lambda V = V\lambda^\sigma, \quad \forall \lambda \in W(k). \tag{5.1}$$

Then there is an equivalence of categories between finite commutative groups G over k of p-power order and finite A-modules that we denote by

$$G \mapsto \mathcal{D}(G). \tag{5.2}$$

It has the following properties:

(1) It commutes with base-change. In particular

$$\mathcal{D}(G^{(p)}) = \mathcal{D}(G) \times_{A,\sigma} A. \tag{5.3}$$

(2) Under this correspondence the map induced by $\mathrm{Fr} : G \longrightarrow G^{(p)}$ is the $\sigma^{(-1)}$-linear map $V : \mathcal{D}(G) \longrightarrow \mathcal{D}(G)$. Similarly, the map $\mathrm{Ver} : G \longrightarrow G^{(1/p)}$ induces the σ-linear map $\mathrm{Fr} : \mathcal{D}(G) \longrightarrow \mathcal{D}(G)$.

(3) There is duality:

$$\mathcal{D}(G^t) = \mathrm{Hom}_A(\mathcal{D}(G), A). \tag{5.4}$$

In particular, G is local-local (resp. local-étale, resp. étale-local) if and only if both F and V are nilpotent on $\mathcal{D}(G)$ (resp. V nilpotent and F is an isomorphism on $\mathcal{D}(G)$; resp. F nilpotent and V is an isomorphism on $\mathcal{D}(G)$).

(4) The order of G is p^r where r is the length of $\mathcal{D}(G)$ as a $W(k)$-module.

Example 5.1. The group α_p has the Dieudonné module k, where F, V and p act as zero. I.e., $A/(F, V)$, where (F, V) denote the left ideal $AF + AV$ generated by F and V.

Example 5.2. The group μ_p has the Dieudonné module k, where p and V act as zero and F acts as Frobenius. (We remark again that we take the covariant Dieudonné module). I.e., $A/(V, 1 - F)$.

Example 5.3. The group scheme $\underline{\mathbb{Z}/p\mathbb{Z}}$ has Dieudonné module k with p and F acting as zero and V acting as the inverse of Frobenius. I.e., $A/(F, 1 - V)$.

Example 5.4. The group scheme M of Example (3.10) has Dieudonné module $A/(F^2, V^2, F + V)$, while the group schemes G_1 and G_2 appearing in Example (3.11) have Dieudonné modules $A/(V, F^2)$ and $A/(F, V^2)$ respectively.

Dieudonné modules are a very powerful tool in studying p-power finite commutative group schemes. For example, the k-forms of a group G over k^{sep} that can be defined over k are given by $H^1(Gal(k^{sep}/k), \mathrm{Aut}(\mathcal{D}(G)))$. This is often readily computable.

EXAMPLE 5.5. Let $G = \alpha_p$ over k^{sep}. Then $\mathrm{End}(\mathcal{D}(\alpha_p)) = k^{sep}$. By Hilbert's 90 we conclude that α_p has no forms!

EXAMPLE 5.6. Let $G = \mu_p$. Then $\mathrm{End}(\mathcal{D}(\mu_p)) = \{\lambda \in k^{sep} : \lambda^p = \lambda\}$. That is $\mathrm{End}(\mathcal{D}(\mu_p)) = \mathbb{F}_p$. Now $H^1(Gal(k^{sep}/k), \mathbb{F}_p^\times) \cong k^\times/(k^\times)^{p-1}$ by Kummer theory.

EXAMPLE 5.7. Consider M. Its Dieudonné module is $k^{sep}e_1 \oplus k^{sep}e_2$ with F acting by $Fe_1 = e_2$, $Fe_2 = 0$ and $Ve_1 = -e_2$, $Ve_2 = 0$. Note that this module is cyclic with generator e_1. An endomorphism f is thus completely determined by $f(e_1) = ae_1 + be_2$. In fact $f(e_2) = f(Fe_1) = Ff(e_1) = a^\sigma e_2$. The conditions on f being a map of Dieudonné modules is that f commutes with V and F. That is $-a^{\sigma^{-1}}e_2 = Vf(e_1) = f(Ve_1) = f(-e_2) = -a^\sigma e_2$. That is, $a \in \mathbb{F}_{p^2}$.

If we identify f with the couple (a, b), then

$$\mathrm{End}(M) = \{(a, b) : a \in \mathbb{F}_{p^2}, b \in k^{sep}\} \tag{5.5}$$

with component-wise addition and multiplication given by

$$(\alpha, \beta)(a, b) = (\alpha a, \beta a + b\alpha^\sigma). \tag{5.6}$$

The identity is $(1, 0)$ and hence

$$\mathrm{Aut}(M) = \{(a, b) : a \in \mathbb{F}_{p^2}^\times, b \in k^{sep}\}. \tag{5.7}$$

One has an exact sequence of $Gal(k^{sep}/k)$ modules:

$$1 \longrightarrow (k^{sep}, +) \longrightarrow \mathrm{Aut}(M) \longrightarrow \mathbb{F}_{p^2}^\times \longrightarrow 1. \tag{5.8}$$

The maps are $\beta \mapsto (1, \beta)$ and $(\alpha, \beta) \mapsto \alpha$. In fact it splits, as Galois modules, by $\alpha \mapsto (\alpha, 0)$. That is, $\mathrm{Aut}(M) = k^{sep} \rtimes \mathbb{F}_{p^2}^\times$.

Let us assume that $\mathbb{F}_{p^2} \subset k$. Since $H^1(Gal(k^{sep}/k), k^{sep}) = 1$ ($(k^{sep}, +)$ is cohomologically trivial), we get an injection

$$H^1(Gal(k^{sep}/k), M) \hookrightarrow H^1(Gal(k^{sep}/k), \mathbb{F}_{p^2}) = k^\times/(k^\times)^{p^2-1}, \tag{5.9}$$

by Kummer theory.

It is easy to see that $H^1(Gal(k^{sep}/k), M) \hookrightarrow H^1(Gal(k^{sep}/k), \mathbb{F}_{p^2})$ is surjective. That would be obvious if we have dealt with usual (abelian) cohomology because $H^2(Gal(k^{sep}/k), k) = 0$. But, directly: Given a cocycle $\beta \in H^1(Gal(k^{sep}/k), \mathbb{F}_{p^2})$, $\sigma \mapsto \alpha(\sigma)$ one lifts it by $\sigma \mapsto (\alpha(\sigma), 0))$.

We now look again at the case of an abelian variety A/k of dimension g. Since k is a perfect field the Verschiebung morphism $A \longrightarrow A^{(1/p)}$ is well defined. Let $A[\mathrm{Fr}]$ and $A[\mathrm{Ver}]$ be the kernel of Frobenius and the kernel of Verschiebung respectively. These are subgroup schemes of A of order p^g. We have an exact sequence:

$$0 \longrightarrow A[\mathrm{Ver}] \longrightarrow A[p] \xrightarrow{\mathrm{Ver}} A[\mathrm{Fr}] \longrightarrow 0. \tag{5.10}$$

We let $\mathcal{D}$ denote the Dieudonné module $\mathcal{D}(A[p])$, and we apply the Dieudonné functor:

$$0 \longrightarrow \mathcal{D}[Fr] \longrightarrow \mathcal{D} \xrightarrow{\mathrm{Fr}} \mathrm{Fr}\mathcal{D} \longrightarrow 0. \tag{5.11}$$

Now, assuming that A has a polarization prime to p, there is an isomorphism of $k[\mathrm{Fr}, \mathrm{Ver}]$-modules of $\mathcal{D}$ with $H^1_{dR}(A)$ (the action on the latter comes from

$H^1_{dR}(A) \cong H^1_{Crys}(A/W(k))/pH^1_{Crys}(A/W(k)))$. [1] The vector space H^1_{dR} has a canonical filtration

$$0 \longrightarrow H^0(A,\Omega^1_A) \longrightarrow H^1_{dR}(A) \longrightarrow H^1(A,\mathcal{O}_A) \longrightarrow 0. \tag{5.12}$$

The arrows are defined by a spectral sequence. The sequences (5.10) and (5.12) are closely related. Indeed, since Fr acts as zero on differential forms, the identification $\mathcal{D} = H^1_{dR}(A)$ implies that $\mathcal{D}[\mathrm{Fr}] = H^0(A,\Omega^1_A)$. We may further identify $H^1(A,\mathcal{O}_A)$ with $\mathrm{Fr}\mathcal{D}$. On the other hand, the map $\mathrm{Fr}:\mathcal{D}\longrightarrow\mathcal{D}$ identifies $\mathcal{D}/\mathcal{D}[F]$ with $\mathcal{D}[V]$, only that its not linear: $\mathrm{Fr}(\lambda a) = \lambda^p \mathrm{Fr}(v)$. We obtain: $H^1(A,O_A) \cong \mathcal{D}[V]\otimes_{k,\mathrm{Fr}} k$. Thus,

$$\begin{array}{ccccccccc} 0 & \longrightarrow & H^0(A,\Omega^1_A) & \longrightarrow & H^1_{dR}(A) & \longrightarrow & H^1(A,\mathcal{O}_A) & \longrightarrow & 0 \\ & & \downarrow\cong & & \downarrow\cong & & \downarrow\cong & & \\ 0 & \longrightarrow & \mathcal{D}(A[V]) & \longrightarrow & \mathcal{D}(A[p]) & \xrightarrow{\mathrm{Fr}} & \mathcal{D}(A[F])\otimes_{k,\mathrm{Fr}} k & \longrightarrow & 0 \end{array} \tag{5.13}$$

We remark that we are working with the covariant Dieudonné module. Oda [**87**] works with the contravariant Dieudonné theory $\mathbb{D}$ and obtains in l.c. Corollary 5.11 an identification

$$\begin{array}{ccccccccc} 0 & \longrightarrow & H^0(A,\Omega^1_A) & \longrightarrow & H^1_{dR}(A) & \longrightarrow & H^1(A,\mathcal{O}_A) & \longrightarrow & 0 \\ & & \downarrow\cong & & \downarrow\cong & & \downarrow\cong & & \\ 0 & \longrightarrow & (k,\mathrm{Fr}^{-1})\otimes_k \mathbb{D}(A[F]) & \longrightarrow & \mathbb{D}(A[p]) & \xrightarrow{\mathrm{Fr}} & \mathbb{D}(A[V]) & \longrightarrow & 0 \end{array} \tag{5.14}$$

6. The α-group

In this section we define the alpha group of an abelian variety $A\longrightarrow S$ over a scheme S of characteristic p. A caveat is that there is no good way to define this subgroup scheme as a group scheme of the abelian variety itself. We therefore define it as a subgroup scheme of the base change $A^{(p)}\longrightarrow S$.

Definition 6.1. Let $A\longrightarrow S$ be an abelian scheme. Let

$$\alpha(A) = \mathrm{Ker}(\mathrm{Ver}: A^{(p)}\longrightarrow A)\cap \mathrm{Ker}(\mathrm{Fr}: A^{(p)}\longrightarrow A^{(p^2)}). \tag{6.1}$$

We call it the alpha group of A.

Some remarks are in order: First, note that $\alpha(A)$ is a group scheme over S. Second, the construction of $\alpha(A)$ is stable under base change: That is, for every morphism $T\longrightarrow S$ we have

$$\alpha(A/S)\times_S T = \alpha(A\times_S T). \tag{6.2}$$

This is nothing more then the behaviour of Ver and Fr under base change and that Ker(Ver) and Ker(Fr) represent the functors "the kernel of Verschiebung" and the "kernel of Frobenius", respectively. Third, let S be a perfect scheme, namely, the absolute Frobenius morphism $\mathrm{Fr}^{abs}: S\longrightarrow S$ is an isomorphism (e.g. S is the

[1] The canonical isomorphism is of the contravariant Dieudonné module $\mathbb{D}(A[p]) = \mathcal{D}(A[p])^t$ with $H^1_{dR}(A)$. However, $\mathcal{D}(A[p])^t = \mathcal{D}(A^t[p])$, which may be identified with $\mathcal{D}(A[p])$ as Dieudonné modules if we have a polarization of degree prime to p.

spectrum of an algebraically closed field). We may therefore write $A = B^{(p)}$ for some abelian scheme $B \longrightarrow S$. Then

$$\alpha(B)^{(p)} = \alpha(A). \tag{6.3}$$

That is, the group scheme $\alpha(A) \subset A^{(p)}$ descends to a group scheme of A. Fourth, as a sheaf in the fppf topology, $\alpha(A)$ is a "constructible sheaf".

DEFINITION 6.2. We say that $A \longrightarrow S$ has a-number greater or equal to a and write

$$a(A) \geq a, \tag{6.4}$$

if the rank of $\alpha(A)$ is greater or equal to a.

In particular, for A over an algebraically closed field k, the a number of A in the sense above, and the a-number of A in the sense of

$$\dim_k(\alpha_p, A), \tag{6.5}$$

are the same. Indeed, first

$$\dim_k(\alpha_p, A) = \dim_k(\alpha_p, A^{(p)}) = \dim_k(\alpha_p, \alpha(A)). \tag{6.6}$$

Secondly, a commutative finite flat group scheme $G \longrightarrow S$ killed by Frobenius and Verschiebung (in the sense that $\mathrm{Ver} : G^{(p)} \longrightarrow G$ is the zero morphism) is locally on S isomorphic to α_p^r/S. Thirdly, over k the embeddings

$$\alpha_p \longrightarrow \alpha_p^r \tag{6.7}$$

are parameterized by surjective maps of Hopf k-algebras

$$k[x_1, \ldots, x_r]/(x_1^p, \ldots, x_r^p) \longrightarrow k[t]/(t^p). \tag{6.8}$$

Consider first that case $r = 1$. Then x_1 is mapped to $f(t) = a_0 + a_1 t + \cdots + a_{p-1}t^{(p-1)}$. Then $x_1 \otimes 1 + 1 \otimes x_1$ is mapped to $f(t) \otimes 1 + 1 \otimes f(t)$ which should be equal to $m^* f(t)$, where $m : \alpha_p \times \alpha_p \longrightarrow \alpha_p$ is the multiplication morphism. But

$$m^* f(t) = a_0 \otimes 1 + a_1(t \otimes 1 + 1 \otimes t) + \ldots a_{p-1}(t \otimes 1 + 1 \otimes t)^{p-1}. \tag{6.9}$$

Equating coefficients we get $f(t) = at$ for some $a \in k$. Thus, coming back to the general case, giving a morphism $\alpha_p \longrightarrow \alpha_p^r$ is equivalent to giving a vector $(a_1, \ldots, a_r) \in k^r$. The correspondence being given by associating to the vector $(a_1, \ldots, a_r)$ the unique morphism $k[x_1, \ldots, x_r]/(x_1^p, \ldots, x_r^p) \longrightarrow k[t]/(t^p)$ taking x_i to $a_i t$. The morphism of groups is injective, if and only of the morphism of algebras is surjective, if and only if $(a_1, \ldots, a_r)$ is not the zero vector. Furthermore, taking the case $r = 1$ we see that the isomorphisms of α_p are in natural bijection with $k^\times$, and the natural action of the automorphism associated to $a \in k^\times$ on the embeddings $\alpha_p \longrightarrow \alpha_p^r$ is given by $(a_1, \ldots, a_r) \mapsto (aa_1, \ldots, aa_r)$. That is, the homomorphisms of α_p to α_p^r are naturally isomorphic to $\mathbb{A}_k^r$, and the subgroups of α_p^r that are isomorphic to α_p are in natural bijection with $\mathbb{P}_k^{(r-1)}$.

Bibliography

[1] Ash, A., Stevens, G.: Cohomology of arithmetic groups and congruences between systems of Hecke eigenvalues. *J. Reine Angew. Math.* 365 (1986), 192–220.

[2] Ash, A., Stevens, G.: Modular forms in characteristic l and special values of their L-functions. *Duke Math. J.* 53 (1986), no. 3, 849–868.

[3] Bachmat, E., Goren, E. Z.: On the non-ordinary locus in Hilbert-Blumenthal surfaces. *Math. Ann.* 313 (1999), no. 3, 475–506.

[4] Berthelot, P., Breen, L., Messing, W.: *Théorie de Dieudonné cristalline. II.* Lecture Notes in Mathematics, 930. Springer-Verlag, 1982.

[5] Borel, A.: *Linear Algebraic Groups*, Second edition, Graduate Texts in Mathematics 126, Springer-Verlag, 1991.

[6] Bosch, S., Lütkebohmert, W., Raynaud, M.: *Néron models.* Ergebnisse der Mathematik und ihrer Grenzgebiete (3), 21. Springer-Verlag, 1990.

[7] Cassou-Noguès, P.: Valeurs aux entiers négatifs des fonctions zêta et fonctions zêta p-adiques. *Invent. Math.* 51 (1979), no. 1, 29–59.

[8] Chai, C.-L.: Every ordinary symplectic isogeny class in positive characteristic is dense in the moduli. *Invent. Math.* 121 (1995), no. 3, 439–479.

[9] Chai, C.-L.: Arithmetic minimal compactification of the Hilbert-Blumenthal moduli spaces. *Ann. of Math.* (2) 131 (1990), no. 3, 541–554.

[10] Chai, C.-L., Norman, P.: Singularities of the $\Gamma_0(p)$-level structure. *J. Algebraic Geometry,* 1 (1992), no.2, 251-278.

[11] Chai, C.-L., Norman, P.: Bad reduction of the Siegel moduli scheme of genus two with $\Gamma_0(p)$-level structure. *Amer. J. Math.*, 112 (1990), no. 6, 1003-1071.

[12] Chevalley C.: Une démonstration d'un théorème sur les groupes algébriques. *J. Mathématiques Pures et Appliquées,* 39 (4), 1960, pp. 307-317.

[13] Cohen, H.: Variations sur un thème de Siegel-Hecke. *Séminaire Delange-Pisot-Poitou (15 année: 1973/74), Théorie des nombres,* Fasc. 1, Exp. No. 14, 7 pp. Secrératariat Mathématique, Paris, 1975.

[14] Coleman, Robert F.: Classical and overconvergent modular forms. *Invent. Math.* 124 (1996), no. 1-3, 215–241.

[15] Conrad, B. : A modern proof of Chevalley's theorem on algebraic groups. unpublished, http:// www-math.mit.edu// dejong// # brian

[16] Cornell G., Silverman, J., editors: *Arithmetic Geometry*, Springer-Verlag, 1986.

[17] Deligne, P.: La conjecture de Weil. I. *Inst. Hautes Études Sci. Publ. Math.* No. 43, (1974), 273–307.

[18] Deligne, P.: La conjecture de Weil. II. *Inst. Hautes Études Sci. Publ. Math.* No. 52, (1980), 137–252.

[19] Deligne, P.: Variétés abéliennes ordinaires sur un corps fini. *Invent. Math.* 8 (1969), 238 - 243.

[20] Deligne, P., Mumford, D.: The irreducibility of the space of curves of given genus. *Inst. Hautes Études Sci. Publ. Math.* No. 36 (1969) 75–109.

[21] Deligne, P., Pappas, G.: Singularités des espaces de modules de Hilbert, en les caractéristiques divisant le discriminant. *Compositio Math.* 90 (1994), no. 1, 59–79.

[22] Deligne, P., Rapoport, M.: Les schémas de modules de courbes elliptiques. *In* Modular functions of one variable, II (Proc. Internat. Summer School, Univ. Antwerp, Antwerp, 1972), pp. 143–316. *Lecture Notes in Mathematics* Vol. 349, Springer, Berlin, 1973.

[23] Deligne, P., Ribet, K. A.: Values of abelian L-functions at negative integers over totally real fields. *Invent. Math.* 59 (1980), no. 3, 227–286.

[24] Demazure, M.: *Lectures on p-divisible groups*. Lecture Notes in Mathematics, Vol. 302. Springer-Verlag, 1972.
[25] Deninger, C.; Murre, J.: Motivic decomposition of abelian schemes and the Fourier transform. *J. Reine Angew. Math.* 422 (1991), 201-219.
[26] de Shalit, Ehud: *Iwasawa theory of elliptic curves with complex multiplication*. Perspectives in Mathematics, 3. Academic Press, 1987.
[27] Diamond, F., Im, J.: Modular forms and modular curves. *In* Seminar on Fermat's Last Theorem (Toronto, ON, 1993–1994), 39–133, *CMS Conf. Proc.*, 17, Amer. Math. Soc., Providence, RI, 1995.
[28] Drinfeld, V. G.: Elliptic modules. *Math. USSR Sbornik*, Vol. 23 (1974), No. 4.
[29] Dwork, B.: The U_p operator of Atkin on modular functions of level 2 with growth conditions. *In* Modular functions of one variable, III (Proc. Internat. Summer School, Univ. Antwerp, Antwerp, 1972), pp. 57–67. *Lecture Notes in Mathematics*, Vol. 350, Springer, 1973.
[30] Ellenberg, J. S.: *Hilbert modular forms and the Galois representations associated to Hilbert-Blumenthal varieties* , Thesis, Harvard University, May 1998.
[31] Faltings,G., Chai,C.-L.: *Degeneration of Abelian Varieties*, Erbegnisse der Mathematik und ihrer Grenzgebiete (3), 22, Springer-Verlag, 1990.
[32] Faltings, G., Wüstholz, G., Grunewald, F., Schappacher, N., Stuhler, U.: *Rational points*. Third edition. Papers from the seminar held at the Max-Planck-Institut für Mathematik, Bonn/Wuppertal, 1983/1984. With an appendix by Wüstholz. Aspects of Mathematics, E6. Friedr. Vieweg & Sohn, 1992.
[33] Farkas, H. M., Kra, I.: *Riemann surfaces*. Second edition. Graduate Texts in Mathematics, 71. Springer-Verlag, 1992.
[34] Fröhlich, A., Taylor, M. J.: *Algebraic number theory* , Cambridge Studies in Advanced Mathematics, 27. Cambridge University Press, Cambridge, 1993.
[35] Demazure, M., Gabriel, P.: *Groupes algébriques. Tome I: Géométrie algébrique, généralités, groupes commutatifs. Avec un appendice* Corps de classes local *par M. Hazewinkel.* Masson & Cie, éditeur, Paris; North-Holland Publishing Co., Amsterdam, 1970.
[36] van der Geer, G. : *Hilbert modular surfaces*. Ergebnisse der Mathematik und ihrer Grenzgebiete (3), 16. Springer-Verlag, 1988.
[37] Hasse invariants for Hilbert modular varieties, *Israel J. Math.*, to appear.
[38] Goren, E. Z.: Hilbert modular forms modulo p^m – the unramified case, CICMA pre-print 1998-10, submitted, 22 pp.
[39] Goren, E. Z., Oort, F.: Stratifications of Hilbert modular varieties, *J. Algebraic Geometry* 9 (2000), 111-154.
[40] Gouvêa, F.Q. : *Deformations of Galois Representations*, P.C.M.I. Lecture Notes.
[41] Griffiths,P., Harris,J. : *Principles of Algebraic Geometry*, John Wiley & Sons, New York, 1978.
[42] Gross, Benedict H.: Heights and the special values of L-series. *In* Number theory (Montreal, Que., 1985), 115–187, *CMS Conf. Proc.*, 7, Amer. Math. Soc., Providence, RI, 1987.
[43] Gross, Benedict H.: A tameness criterion for Galois representations associated to modular forms (mod p). *Duke Math. J.* 61 (1990), no. 2, 445–517.
[44] Grothendieck, A.: Eléments de géométrie algébrique. III. Étude cohomologique des faisceaux cohérents. I. *Inst. Hautes Études Sci. Publ. Math.* No. 11 (1961).
[45] Grothendieck, A.: Eléments de géométrie algébrique. III. Étude cohomologique des faisceaux cohérents. II. *Inst. Hautes Études Sci. Publ. Math.* No. 17 (1963).
[46] Grothendieck, A.: Schémas en groupes. I, II, III: Séminaire de Géométrie Algébrique du Bois Marie 1962/64 (SGA 3). Dirigé par M. Demazure et A. Grothendieck. *Lecture Notes in Mathematics*, Vol. 151, 152, 153, Springer-Verlag, Berlin-New York 1962/1964
[47] Hartshorne, R.: *Algebraic geometry*, Graduate Texts in Mathematics 52, Springer-Verlag 1977.
[48] Hazewinkel, M.: *Formal groups and applications*. Pure and Applied Mathematics, 78. Academic Press, 1978.
[49] Honda, T.: On the theory of commutative formal groups. *J. Math. Soc. Japan* 22 (1970) 213–246.
[50] Igusa, J.-I.: Class number of a definite quaternion with prime discriminant. *Proc. Nat. Acad. Sci. U.S.A.* 44 1958 312–314.

[51] Ireland, K. F., Rosen, M. I.: *A classical introduction to modern number theory*. Graduate Texts in Mathematics, 84. Springer-Verlag, 1982.
[52] Iwasawa K. : *Lectures on p-adic L-functions*, Annals of Mathematics Studies, No. 74. Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, 1972.
[53] de Jong, A.J.: *Moduli of abelian varieties and Dieudonné modules of finite group schemes*. Ph.D. Thesis, Rijksuniversiteit Utrecht, 1992.
[54] de Jong, A. J., Oort, F.: Purity of the stratification by Newton polygons. *J. Amer. Math. Soc.* 13 (2000), no. 1, 209–241.
[55] Jochnowitz, N.: Congruences between systems of eigenvalues of modular forms. *Trans. Amer. Math. Soc.* 270 (1982), no. 1, 269–285.
[56] Jochnowitz, N.: A study of the local components of the Hecke algebra mod l. *Trans. Amer. Math. Soc.* 270 (1982), no. 1, 253–267.
[57] Jochnowitz, N.: The index of the Hecke ring, T_k, in the ring of integers of $T_k \otimes Q$. *Duke Math. J.* 46 (1979), no. 4, 861–869.
[58] Katz, N. M.: p-adic properties of modular schemes and modular forms. *In* Modular functions of one variable, III (Proc. Internat. Summer School, Univ. Antwerp, Antwerp, 1972), pp. 69–190. *Lecture Notes in Mathematics*, Vol. 350, Springer-Verlag, 1973.
[59] Katz, N. M.: Higher congruences between modular forms. *Ann. of Math.* (2) 101 (1975), 332–367.
[60] Katz, N. M.: p-adic L-functions for CM fields. *Invent. Math.* 49 (1978), no. 3, 199–297.
[61] Katz, N. M.: Serre-Tate local moduli. *In* Algebraic surfaces (Orsay, 1976–78), pp. 138–202, *Lecture Notes in Mathematics*, 868, Springer, 1981.
[62] Katz, N. M.: Slope filtration of F-crystals. *In* Journées de Géométrie Algébrique de Rennes (Rennes, 1978), Vol. I, pp. 113–163, *Astérisque*, 63, Soc. Math. France, Paris, 1979.
[63] Kodaira, K.; Spencer, D. C.: On deformations of complex analytic structures. I, II. *Ann. of Math.* (2) 67 (1958), 328–466.
[64] Kolster, M., Nguyen Quang Do, T.: Syntomic regulators and special values of p-adic L-functions. *Invent. Math.* 133 (1998), no. 2, 417–447.
[65] Lang, S.: *Complex multiplication*. Grundlehren der Mathematischen Wissenschaften 255. Springer-Verlag, 1983.
[66] Lang, Serge: *Elliptic functions*. With an appendix by J. Tate. Second edition. Graduate Texts in Mathematics, 112. Springer-Verlag, 1987.
[67] Lange H., Birkenhake, Ch.: *Complex Abelian Varieities*, Grundlehren der mathematischen Wissenschaften 302, Springer-Verlag 1992.
[Laz] Lazard, M.: *Commutative formal groups*. Lecture Notes in Mathematics, Vol. 443. Springer-Verlag, 1975.
[68] Manin, Yu.I.: The theory of commutative formal groups over fields of finite characteristic. *Russ. Math. Surveys* 18 (1963), 1 - 80.
[69] Messing, W.: *The crystals associated to Barsotti-Tate groups: with applications to abelian schemes*. Lecture Notes in Mathematics, Vol. 264. Springer-Verlag, Berlin-New York, 1972.
[70] Mestre, J.-F.: La méthode des graphes. Exemples et applications. *Proceedings of the international conference on class numbers and fundamental units of algebraic number fields* (Katata, 1986), 217–242, Nagoya Univ., Nagoya, 1986.
[71] Milne, J. S.: *Étale cohomology*. Princeton Mathematical Series, 33. Princeton University Press, Princeton, N.J., 1980.
[72] Milne, J. S.: *Abelian Varieties* , Notes for Math 731, www.math.lsa.umich.edu/ jmilne/.
[73] Milne, J. S.: Jacobian Varieties. In *Arithmetic geometry*. Papers from the conference held at the University of Connecticut, Storrs, Conn., July 30–August 10, 1984. Edited by Gary Cornell and Joseph H. Silverman. Springer-Verlag, 1986.
[74] Milne, J.S. : Points on Shimura varieties mod p. In *Automorphic forms, representations and L-functions* Part 2, pp. 165-184, Proc.Sympos.Pure math., XXXIII, Amer. Math. Soc., Providence, R.I., 1979
[75] Mumford, D. : *Abelian varieties*, Oxford University Press, 1970.
[76] Mumford, D. : On the equations defining abelian varieties I. *Invent. Math.*, 1 (1966) pp. 287-354.
[77] Mumford, D. : Picard groups of moduli problems. In: O. Schilling (Ed.), *Arithmetic Algebraic Geometry*, (Proc. Conf. Purdue Univ., 1963) pp. 33–81 Harper & Row, 1965.

[78] Mumford, D.: *The red book of varieties and schemes*. Lecture Notes in Mathematics, 1358. Springer-Verlag, 1988.

[79] Mumford, D.: Towards an enumerative geometry of the moduli space of curves. In *Arithmetic and geometry*, Vol. II, 271–328, Progr. Math., 36, Birkhäuser Boston, Boston, Mass., 1983.

[80] Mumford, D.: *Tata lectures on theta. I*. With the assistance of C. Musili, M. Nori, E. Previato and M. Stillman. Progress in Mathematics, 28. Birkhäuser Boston, Inc., Boston, Mass., 1983.

[81] Mumford, D.: *Tata lectures on theta. II*. Jacobian theta functions and differential equations. With the collaboration of C. Musili, M. Nori, E. Previato, M. Stillman and H. Umemura. Progress in Mathematics, 43. Birkhäuser Boston, Inc., Boston, MA, 1984.

[82] Mumford, D.: *Tata lectures on theta. III*. With the collaboration of Madhav Nori and Peter Norman. Progress in Mathematics, 97. Birkhäuser Boston, Inc., Boston, MA, 1991. viii+202 pp.

[83] Mumford, D., Fogarty, J., Kirwan, F.: *Geometric invariant theory*. Third edition. Ergebnisse der Mathematik und ihrer Grenzgebiete (2), 34. Springer-Verlag, 1994.

[84] Murre, J. P.: Algebraic cycles on abelian varieties, Application of abstract Fourier theory. In *The Arithmetic and Geometry of Algebraic cycles*, NATO Science series, Series C: Mathematical and Physical Sciences, Vol. 548, Kluwer Academic Publishers 2000, 307 - 320.

[85] Murty, V. K.: *Introduction to abelian varieties*. CRM Monograph series 3. American Mathematical Society, Providence, RI, 1993.

[86] Norman, P. , Oort, F.: Moduli of abelian varieties. *Ann. of Math.* (2) 112 (1980), no. 3, 413–439.

[87] Oda, T.: The first de Rham cohomology group and Dieudonné modules. *Ann. Sci. Ècole Norm. Sup.* (4) 2 (1969) 63–135.

[88] Ogg, A. P.: *Modular forms and Dirichlet series*. W. A. Benjamin, Inc., New York-Amsterdam 1969.

[89] Ogg, A. P.: Rational points on certain elliptic modular curves. In *Analytic number theory* (Proc. Sympos. Pure Math., Vol XXIV, St. Louis Univ., St. Louis, Mo., 1972), pp. 221–231. Amer. Math. Soc., Providence, R.I., 1973.

[90] Ogg, A. P.: Hyperelliptic modular curves. *Bull. Soc. Math. France* 102 (1974), 449–462.

[91] Oort, F.: Finite group schemes, local moduli for abelian varieties, and lifting problems. In: *Algebraic geometry*, Oslo, 1970. Proceedings of the Fifth Nordic Summer School in Mathematics held in Oslo, August 5–25, 1970. Edited by F. Oort. Wolters-Noordhoff Publishing, Groningen, 1972.

[92] Oort, F.: Moduli of abelian varieties and Newton polygons. *C. R. Acad. Sci. Paris* Sér. I Math. 312 (1991), no. 5, 385–389.

[93] Oort, F.: Which abelian surfaces are products of elliptic curves? *Math. Ann.* 214 (1975), 35–47.

[94] Oort F.: *Commutative group schemes*. Lecture Notes in Mathematics, 15, Springer-Verlag, 1966.

[95] Pappas, G.: Arithmetic models for Hilbert modular varieties. *Compositio Math.* 98 (1995), no. 1, 43–76.

[96] Pizer, A.: An algorithm for computing modular forms on $\Gamma_0(N)$. *J. Algebra* 64 (1980), no. 2, 340–390.

[97] Rapoport, M. : Compactifications de l'espace de modules de Hilbret-Blumenthal. *Compositio Math.* 36 (1978), no.3, 255-335.

[98] Ribet, K.: p-adic interpolation via Hilbert modular forms. In *Algebraic geometry* (Proc. Sympos. Pure Math., Vol. 29, Humboldt State Univ., Arcata, Calif., 1974) pp. 581-592, Amer. Math. Soc., Providence, R.I., 1976.

[99] Ribet K., Stein, W.A.: *Lectures on Serre's conjectures*, to appear.

[100] Rück, H.-G.: Abelian surfaces and Jacobian varieties over finite fields. *Compositio Math.* 76 (1990), no.3,351-36.

[101] Serre, J.-P.: Congruences et formes modulaires [d'après H. P. F. Swinnerton-Dyer]. Séminaire Bourbaki, 24e année (1971/1972), Exp. No. 416, pp. 319–338. *Lecture Notes in Mathematics*, Vol. 317, Springer, Berlin, 1973.

[102] Serre, J.-P.: Formes modulaires et fonctions zêta p-adiques. In *Modular functions of one variable, III* (Proc. Internat. Summer School, Univ. Antwerp, 1972), pp. 191–268. Lecture Notes in Mathematics, Vol. 350, Springer, Berlin, 1973.

[103] Serre, J.-P.: Sur les représentations modulaires de degré 2 de Gal($\overline{Q}/Q$). *Duke Math. J.* 54 (1987), no. 1, 179–230.

[104] Serre, J. P.: Groupes p-divisibles (d'après J. Tate), *Séminaire Bourbaki*, Vol. 10, Exp. No. 318, 73–86.

[105] Shimura, G.: On analytic families of polarized abelian varieties and automorphic functions. *Ann. of Math.* (2) 78 (1963), 149–192.

[106] Shimura, G.: The special values of the zeta functions associated with Hilbert modular forms. *Duke Math. J.* 45 (1978), no. 3, 637–679.

[107] Silverman, J. H.: *The arithmetic of elliptic curves*. Graduate Texts in Mathematics, 106. Springer-Verlag, 1986.

[108] Silverman, J. H.: *Advanced topics in the arithmetic of elliptic curves.* Graduate Texts in Mathematics, 151. Springer-Verlag, 1994.

[109] Swinnerton-Dyer, H. P. F.: *Analytic theory of Abelian varieties*. London Mathematical Society lecture note series 14, Cambridge University Press, 1974.

[110] Swinnerton-Dyer, H. P. F.: On l-adic representations and congruences for coefficients of modular forms. In *Modular functions of one variable, III* (Proc. Internat. Summer School, Univ. Antwerp, 1972), pp. 1–55. Lecture Notes in Mathematics, Vol. 350, Springer, 1973.

[111] Tate, J.: On the conjectures of Birch and Swinnerton-Dyer and a geometric analogue. *Séminaire Bourbaki*, Vol.9, Exp. No. 306, 415-440, Soc. Math. France, Paris, 1995.

[112] Tate, J.: Endomorphisms of abelian varieties over finite fields. *Invent. Math.* 2 (1966) 134–144.

[113] Tate, J.: $p-divisible\ groups$. In *Proc. Conf. Local Fields* (Driebergen, 1966) pp. 158–183 Springer-Verlag, 1967.

[114] Tate, J.: Finite flat group schemes. In *Modular forms and Fermat's last theorem* (Boston, MA, 1995), 121–154, Springer, 1997.

[115] Tsuyumine, S.: On values of L-functions of totally real algebraic number fields at integers. *Acta Arith.* 76 (1996), no. 4, 359–392.

[116] Ulmer, D. L.: On universal elliptic curves over Igusa curves. *Invent. Math.* 99 (1990), no. 2, 377–391.

[117] Washington, L. C.: *Introduction to Cyclotomic Fields. Second Edition*, Graduate Text in Mathematics 83, Springer-Verlag, 1997.

[118] Waterhouse, W. C.: *Introduction to affine group schemes*. Graduate Texts in Mathematics, 66. Springer-Verlag, 1979.

[119] Waterhouse, W.C., Milne, J.S.: Abelian varieties over finite fields. In *1969 Number Theory Institue* (Prov. Sympos. Pure Math., Vol. XX, State Univ. New York, Stony Brook, N.Y., 1969),pp.53-64. Amer. Math. Soc, Providence, R.I., 1971.

[120] Wedhorn, T.: Ordinariness in good reductions of Shimura varieties of PEL-type. *Ann. Sci. École Norm. Sup.* (4) 32 (1999), no. 5, 575–618.

[121] Weil, A.: Sur certains groupes d'opérateurs unitaires. *Acta Math.* 111 (1964), 143–211.

[122] Weil, A.: Sur la formule de Siegel dans la théorie des groupes classiques. *Acta Math.* 113 (1965), 1–87.

[123] Yui, N.: On the Jacobian varieties of hyperelliptic curves over fields of characteristic $p > 2$. *J.Algebra* 52 (1978), no. 2, 378-410.

[124] Yui, N.: Formal groups and some arithmetic properties of elliptic curves. In *Algebraic geometry* (Proc. Summer Meeting, Univ. Copenhagen, Copenhagen, 1978), pp. 630–658, Lecture Notes in Mathematics, 732, Springer, Berlin, 1979.

[125] Zarhin, J. G.: Isogenies of abelian varieties over fields of finite characteristics. *Math. USSR Sb.* 24 (1974), 451-461.

[126] Zarhin, J. G.: A remark on endomorphisms of abelian varieties over function fields of finite characteristics. *Math. USSR Isv.* 8 (1974), 477-480.

[127] Zink, Th.: *The display of a formal p-divisible group*. Universität Bielefeld Preprint 98-017, February 1998, 155 pp.

Index