(6867)

J. reine angew. Math. ▮▮ (2011), 1—26
DOI 10.1515/CRELLE.2011.▮▮▮

Journal für die reine und angewandte Mathematik

Higgs bundles over the good reduction of a quaternionic Shimura curve

By *Mao Sheng* at Mainz, *Jiajin Zhang* at Chengdu and *Kang Zuo* at Mainz

Abstract. This paper is devoted to the study of the Higgs bundle associated with the universal abelian variety over the good reduction of a Shimura curve of PEL type. Due to the endomorphism structure, the Higgs bundle decomposes into the direct sum of Higgs subbundles of rank two. They are basically divided into two types: *uniformizing type* and *unitary type*. As the first application we obtain the mass formula counting the number of geometric points of the degeneracy locus in the Newton polygon stratification. We show that each Higgs subbundle is Higgs semistable. Furthermore, for each Higgs subbundle of unitary type, either it is strongly semistable, or its Frobenius pull-back of a suitable power achieves the upper bound of the instability. We describe the Simpson–Ogus–Vologodsky correspondence for the Higgs subbundles in terms of the classical Cartier descent.

1. Introduction

Let D be a quaternion division algebra over a totally real field F which is exactly split at one infinite place of F. By choosing additionally a totally imaginary quadratic field extension K of F, the data (D, K) allows one to define a Shimura curve of PEL type (see [2]). In this paper, we study the Higgs bundle (E, θ) associated with the universal abelian variety over $\mathscr{M}_0$, which is one of the geometrically connected components of the good reduction of this Shimura curve modulo p. The passage of the Higgs bundle from char 0 to char p has two aims. The first is to study the Newton polygon stratification of the moduli space in char p. A similar method has already been extensively employed in recent years (for example, see [9], [6], [21]). The prototype of such study is the supersingular locus in the moduli space of elliptic curves and the classical Deuring formula. From this example one sees a basic phenomenon occurring in the geometry of a moduli space in char p, namely the degeneration of the relative Frobenius morphism along certain algebraic sublocus of the whole moduli space. The second aim is to investigate the relation between the Higgs bundles over a char p (or p-adic) field and the topology of the underlying spaces in a char p (or p-adic) field. In the classical situation, that is, over the field of complex numbers, this is beautifully

This work is supported by the SFB/TR 45 'Periods, Moduli Spaces and Arithmetic of Algebraic Varieties' of the DFG.

expressed in the work of C. Simpson (see [28]). Recently, analogous theories over a char p or p-adic ground field have emerged (see [25], [5] and [8]). We intend to apply these new theories to study the Higgs bundles over Shimura curves of PEL type in char p and mixed characteristic.

Our results are built on the previous work on the Shimura curves of PEL type, particularly the work of Carayol [1] and Deligne [2] and the book by Reimann [26]. Let $\mathscr{M}_0$ be the good reduction in char p of the Shimura curve of PEL type associated with the quaternion division algebra D and the imaginary quadratic field K (see Section 2 for details), and let $f_0 : \mathscr{X}_0 \to \mathscr{M}_0$ be the universal abelian variety and (E, θ) be the associated Higgs bundle. Let g be the genus of the Shimura curve $\mathscr{M}_0$ which is strictly greater than one by our choice of the level structure (see the end of Section 2). Most of the other notations appearing in the following are collected at the end of this section.

Theorem 1.1 (Proposition 4.1 and 4.4). *The Higgs bundle (E, θ) decomposes into the direct sum of rank two Higgs subbundles*:

$$(E, \theta) = \bigoplus_{\phi \in \Phi} (E_\phi, \theta_\phi) \oplus (E_{\bar{\phi}}, \theta_{\bar{\phi}}),$$

where the endomorphism subalgebra $\mathcal{O}_{LK} \subset \mathcal{O}_B$ acts on the summand E_ϕ (resp. $E_{\bar{\phi}}$) via the character $\phi \bmod p$ (resp. $\bar{\phi} \bmod p$). Assume further that $p \geqq 2g$. Then for each $\phi \in \Phi$ (resp. $\bar{\phi} \in \bar{\Phi}$) with $\phi|_F = \tau$ (resp. $\bar{\phi}|_F = \bar{\tau}$), the Higgs subbundle (E_ϕ, θ_ϕ) (resp. $(E_{\bar{\phi}}, \theta_{\bar{\phi}})$) is of maximal Higgs field (see [30]). Each of the remaining Higgs subbundles in the above decomposition is of trivial (or equivalently, zero) Higgs field.

The Higgs subbundles of maximal Higgs field are called of *uniformizing type*; while those of zero Higgs field are called of *unitary type*. In char 0, a Higgs bundle of uniformizing type provides the uniformization of the base Shimura curve (see [30]) while that of unitary type corresponds to a unitary representation of the topological fundamental group. We then analyze the behavior of the iterated Frobenius morphism on the $(0,1)$-component of each Higgs subbundle and derive the following results.

Theorem 1.2 (Corollary 3.3 and Theorem 5.6). *There are only two types of Newton polygons in $\mathscr{M}_0(\mathbb{F})$. Let $\mathscr{S}$ be the jumping locus of the Newton polygons. Then one has the following formula in the Chow ring of $\mathscr{M}_0$*:

$$\mathscr{S} = \frac{1}{2}(1 - p^{[F_\mathfrak{p}:\mathbb{Q}_p]})c_1(\mathscr{M}_0).$$

Taking the degree, one obtains the following mass formula for the Shimura curve $\mathscr{M}_0$:

$$|\mathscr{S}| = (1 - p^{[F_\mathfrak{p}:\mathbb{Q}_p]})(1 - g).$$

From this formula one sees that the number of closed points in the jumping locus of the Newton polygons is proportional to the topological Euler characteristic of the Shimura curve $\mathscr{M}_0$. We would like to make the following conjecture.

Conjecture 1.3. *Let $\mathscr{M}$ be the Shimura curve of Hodge type defined by $G_\mathbb{Q} \to GSp_\mathbb{Q}$ with a suitable level structure, where $G_\mathbb{Q}$ is the $\mathbb{Q}$-group of the units of a quaternion division algebra D over a totally real field F by restriction of scalars. Let p be a prime number such* *that $\mathscr{M}_0$ is the good reduction of $\mathscr{M}$ modulo p. Then there are exactly two possible Newton polygons for the closed points of $\mathscr{M}_0(\mathbb{F})$ and the cardinality of the jumping locus of the Newton polygons is equal to $\frac{1}{2}(1-p^d)\chi_{\mathrm{top}}\big(\mathscr{M}_0(\mathbb{C})\big)$, where d is the local degree $[F_{\mathfrak{p}} : \mathbb{Q}_p]$ and $F_{\mathfrak{p}}$ is a splitting field of D over p.*

In the case of Mumford's families of abelian varieties the work of R. Noot provides strong evidence for the above conjecture. He studied the potential good reduction of a single abelian variety in a Mumford's family, as well as its possible Newton polygons, using Fontaine's theory (see [23] and references therein). From the proof of the mass formula one notices that certain Higgs subbundles of unitary type have no contribution to the jumping of the Newton polygons since the iterated relative Frobenius morphisms do not degenerate on these subbundles. It turns out that these Higgs subbundles of unitary type carry the extra property of being strongly semistable, and this property characterizes the Higgs subbundles of unitary type with non-degenerate iterated relative Frobenius actions. This motivates us to study the (Higgs) stability of the Higgs subbundles under the Frobenius pull-backs in general. For a semistable vector bundle E over a smooth projective curve C in char p, the invariant $v(F_C^*E)$ (see Section 6 for the definition) measures the extent of the instability under the Frobenius pull-back. It is non-negative by definition and it is zero if and only if F_C^*E is still semistable. E is strongly semistable if $v(F_C^{n*}E) = 0$ for all $n \geqq 1$. It is well known that $v(F_C^*E)$ has the upper bound $\big(\mathrm{rank}(E)-1\big)\big(2g(C)-2\big)$ (see Theorem 6.2). Thus the extreme opposite of the strongly semistability is the case that

Note 1: Throughout the paper we wrote "semistability" without hyphen (similar to semistable). Do you agree?

Note 2: $\leqq$ in $n \leqq n_0$ correct?

$$v(F_C^{n*}E) = 0, \quad 1 \leqq n \leqq n_0,$$
$$v(F_C^{n_0*}E) = \big(\mathrm{rank}(E)-1\big)\big(2g(C)-2\big).$$

We write the $\mathrm{Gal}(L\,|\,\mathbb{Q})$-orbit of Φ containing the uniformizing place τ as follows:

$$\mathrm{Hom}_{\mathbb{Q}_p}(L_{\mathfrak{p}}, \overline{\mathbb{Q}}_p) = \{\phi_1, \dots, \phi_d, \phi_1^*, \dots, \phi_d^*\}$$

with $\phi_1|_F = \phi_1^*|_F = \tau$ such that the Frobenius automorphism $\sigma \in \mathrm{Gal}(L_{\mathfrak{p}}\,|\,\mathbb{Q}_p)$ acts on the orbit via the cyclic permutation. Thus we have the following theorem.

Theorem 1.4 (Proposition 6.1, 6.3, and 6.6). *Assume that $p \geqq 2g$. Then the following statements are true*:

(i) *Each Higgs subbundle in Theorem* 1.1 *is Higgs semistable of slope zero. In particular, the Higgs subbundles of unitary type are semistable.*

(ii) *For $\phi \notin \mathrm{Hom}_{\mathbb{Q}_p}(L_{\mathfrak{p}}, \overline{\mathbb{Q}}_p)$, the Higgs subbundle (E_ϕ, θ_ϕ) of unitary type is strongly semistable* (*even étale trivializable*). *The same is true for its bar counterpart.*

(iii) *For $\phi_i \in \mathrm{Hom}_{\mathbb{Q}_p}(L_{\mathfrak{p}}, \overline{\mathbb{Q}}_p)$ with $i \neq 1$, the Higgs subbundle $(E_{\phi_i}, \theta_{\phi_i})$ of unitary type satisfies*

$$v(F_{\mathscr{M}_0}^{n*}E_{\phi_i}) = 0, \quad 1 \leqq n \leqq d-i,$$
$$v(F_{\mathscr{M}_0}^{d-i+1*}E_{\phi_i}) = 2g-2.$$

The same is true for the ϕ_i^*-summand with $i \neq 1$ and for its bar counterpart.

For several reasons we are motivated to examine the Simpson–Ogus–Vologodsky correspondence (see [25]) for the Higgs subbundles in Theorem 1.1. Let $(\mathscr{H}^1_{dR}, \nabla)$ be the first relative de Rham bundle of f_0 with the canonical Gauss–Manin connection, which has also an eigen-decomposition under the $\mathcal{O}_{LK}$-action (see Proposition 4.1). Let F_{con} be the conjugate filtration on $\mathscr{H}^1_{dR}$, which is equally as important as the Hodge filtration in char p geometry.

Theorem 1.5 (Theorem 7.3 and Corollary 7.4). *Assume that*

$$p \geqq \max\{2g, 2([F:\mathbb{Q}]+1)\}.$$

Then for each $\phi \in \Phi$, the Cartier transform of the direct summand $(\mathscr{H}^1_{dR,\phi}, \nabla_\phi)$ of $(\mathscr{H}^1_{dR}, \nabla)$ is just the Cartier descent of $Gr_{F_{\mathrm{con}}}(\mathscr{H}^1_{dR,\phi}, \nabla_\phi)$. The same is true for its bar counterpart. As a consequence, for ϕ_i in Theorem 1.4(iii), *the Harder–Narasimhan filtration on $F^{d-i+1*}_{\mathscr{M}_0} E_{\phi_i}$ is identified with the Hodge filtration on $\mathscr{H}^1_{dR,\phi_1^*}$. It is similar for the star and bar counterparts.*

By the above theorem one sees that the non-strongly semistable Higgs subbundles of unitary type are closely related to the Higgs subbundles of uniformizing type. In some sense one should consider these two types of Higgs subbundles as the same one. Compared with its char 0 analogue, the topological meaning of the Higgs subbundles of uniformizing type is still unclear to us.

The paper is organized as follows. In Section 2 the construction of a Shimura curve of PEL type is briefly reviewed. In Section 3 some known results about Dieudonné modules of the abelian varieties corresponding to the points on $\mathscr{M}_0(\mathbb{F})$ are summarized. In Section 4 the decomposition of the Higgs bundle and the basic properties of the Higgs subbundles are established. Applying the results in Section 3 and Section 4, we obtain the mass formula for the Shimura curve $\mathscr{M}_0$ in Section 5. In Section 6 the Higgs semistability as well as the semistability under Frobenius pull-backs of the Higgs subbundles are discussed. The description of the Simpson–Ogus–Vologodsky correspondence for the Higgs subbundles is contained in Section 7.

Notations and Conventions. (i) For a prime $\mathfrak{q}$ of a number field E, $E_{\mathfrak{q}}$ means the completion of E with respect to $\mathfrak{q}$. For a field E of char 0 (local or global), $\mathcal{O}_E$ is the ring of integers in E and $\bar{E}$ is an algebraic closure of E. $\mathbb{Q}_p^{ur}$ is the maximal unramified subextension of $\mathbb{Q}_p$. Denote by k a finite field of char p and by $\mathbb{F}$ an algebraic closure of k. Let $\sigma \in \mathrm{Gal}(\bar{\mathbb{F}} \,|\, \mathbb{F}_p)$ be the Frobenius automorphism, defined by $x \mapsto x^p$. It is restricted to the Frobenius automorphism of k. For $\mathbb{F}$, denote by $W(\mathbb{F})$ the ring of Witt vectors and one has the canonical lifting of the Frobenius automorphism of $\mathbb{F}$ to $W(\mathbb{F})$, which is again denoted by σ. It is similar for $W(k)$ and k.

(ii) In this paper, F is a fixed totally real number field of degree $n \geqq 2$, and p is a rational prime number which is unramified in F. K and L are two fixed imaginary quadratic field extensions of F (see Section 2 for details). We put $\Psi = \mathrm{Hom}_{\mathbb{Q}}(F, \mathbb{R})$ and $\Phi = \mathrm{Hom}_{\mathbb{Q}}(L, \bar{\mathbb{Q}})$. D is a fixed quaternion division algebra over F, which is exactly split at one infinite place $\tau \in \Psi$ of F.

(iii) For an algebraic variety X over k, one denotes by F_X the absolute Frobenius morphism. For a morphism $f: X \to Y$ over k one has the following commutative diagram of Frobenius morphisms:

$$\begin{tikzcd} X \arrow[r, "\mathscr{F}_{X|Y}"] \arrow[dr, "f"'] & X' \arrow[r, "\pi_{X|Y}"] \arrow[d, "f'"] & X \arrow[d, "f"] \\ & Y \arrow[r, "F_Y"'] & Y. \end{tikzcd}$$

where the square in the diagram is the fiber product, $\mathscr{F}_{X|Y}$ is the relative Frobenius morphism and $\pi_{X|Y} \circ \mathscr{F}_{X|Y} = F_X$. For a vector bundle $\mathscr{E}$ over X in char p, sometimes we denote the (iterated) Frobenius pull-back $F_X^{*n}\mathscr{E}$ $(n \geqq 1)$ by $\mathscr{E}^{(p^n)}$.

(iv) In this paper, the term '*reduction modulo p*' means the following: let R be a DVR of mixed characteristic $(0, p)$ with the residue field $k(R)$, and M be an object defined over R, which can be a module or a scheme. Then the reduction of M modulo p is the base change of M from R to $k(R)$.

Acknowledgment. We would like to thank the referee for his/her careful reading of our paper and helpful advice. We thank C. Deninger for useful discussions on [4] and [5]. Special thanks go to A. Langer for his useful comments on Section 6 and particularly the clarification of a main result in [17] (see Theorem 6.5).

2. Quaternion division algebras and the good reduction of a Shimura curve

Let D be a quaternion division algebra over F, which is split at the infinite place τ and ramified at all remaining infinite places. That is, one has the following isomorphisms:

$$D \otimes_{F,\tau} \mathbb{R} \cong M_2(\mathbb{R}), \quad \text{and} \quad D \otimes_{F,\psi} \mathbb{R} \cong \mathbb{H} \quad \text{for } \psi \neq \tau,$$

where $\mathbb{H}$ is the Hamiltonian quaternion algebra over $\mathbb{R}$. One considers the F-group of the units $D^\times$ as a $\mathbb{Q}$-group by restriction of scalars and defines a homomorphism of real algebraic groups:

$$\begin{aligned} h_D : \mathbb{S} = \mathbb{C}^\times &\to D^\times(\mathbb{R}) \cong \mathrm{GL}_2(\mathbb{R}) \times (\mathbb{H}^\times)^{n-1}, \\ z = x + iy &\mapsto (\begin{pmatrix} x & y \\ -y & x \end{pmatrix}, 1, \dots, 1). \end{aligned}$$

The $D^\times(\mathbb{R})$-conjugacy class X of h_D defines a Shimura curve Sh_D over the reflex field F, where F is considered as a subfield of $\mathbb{C}$ via the embedding τ. For every open compact subgroup $C \subset D^\times(\mathbb{A}^f)$, $Sh_{D,C} = Sh_D/C$ is a projective curve over F, and one has the identification of its complex points

$$Sh_{D,C}(\mathbb{C}) = D^\times(\mathbb{Q}) \backslash \big(X \times D^\times(\mathbb{A}^f)\big)/C,$$

where $\mathbb{A}_f$ is the ring of finite adèles of $\mathbb{Q}$ and $D^\times$ acts on X by the conjugation and on the second summand by the left multiplication.

In certain cases $Sh_{D,C}(\mathbb{C})$ is known to parameterize the principally polarized abelian varieties over $\mathbb{C}$ with special Mumford–Tate groups (see for example [22], §4 and [30], §5). It belongs to the category of Shimura varieties of Hodge type. In this paper we are going to study a related Shimura curve Sh_G which is of PEL type. However the Shimura curve of Hodge type provides the motivation for the further study of the current paper. Now we recall the so-called 'modèle étrange' construction in [2]. First we need to choose an imaginary quadratic field $\mathbb{Q}(\alpha)$ with $\alpha \in \mathbb{C}$ such that p is split in it. We put the composite $K = F(\alpha)$ and it will be fixed in the whole discussion. Considering $F^\times$ and $K^\times$ as $\mathbb{Q}$-groups, we define a new $\mathbb{Q}$-group G by the following short exact sequence:

$$1 \to F^\times \to D^\times \times K^\times \xrightarrow{\pi} G \to 1,$$

where $F^\times \to D^\times \times K^\times$ is given by $f \mapsto (f, f^{-1})$. We fix a subset $\Psi_K \subset \mathrm{Hom}_{\mathbb{Q}}(K, \mathbb{C})$ which induces a bijection to Ψ by restriction to F. Note that Ψ_K is obtained by the trivial extensions of all embeddings of F into $\mathbb{C}$ to embeddings of $K = F + F\alpha$ into $\mathbb{C}$. One has an identification $K^\times(\mathbb{R}) \cong \prod_{\psi \in \Psi} \mathbb{C}^\times$ and defines

$$h_K : \mathbb{S} \to K^\times(\mathbb{R}) = \mathbb{C}^\times \times \prod_{\psi \neq \tau} \mathbb{C}^\times, \quad z \mapsto (1, z, \ldots, z).$$

Let X' denote the conjugacy class of

$$h_G = \pi_{\mathbb{R}} \circ (h_D \times h_K) : \mathbb{S} \to G(\mathbb{R}).$$

It defines a Shimura curve Sh_G over K, where K is a subfield of $\mathbb{C}$ via the map $\tau \in \Psi \cong \Psi_K$. A compact open subgroup C of $G(\mathbb{A}_f)$ defines a projective curve $Sh_{G,C}$. For a suitable $C' \subset D^\times(\mathbb{A}_f)$, the neutral component of $Sh_{D,C'}$ and $Sh_{G,C}$ are isomorphic to each other over certain number field (see [26], §1). The Shimura curve $Sh_{G,C}$ parameterizes the isogeny classes of abelian varieties over K with PEL structure which we describe briefly as follows. Let $B = D \otimes_F K$. Define the natural involution on B by the formula

$$(x \otimes y)' = x^* \otimes \bar{y},$$

where $*$ is the main involution on D and $^-$ is the complex conjugation on K which is the generator of $\mathrm{Gal}(K \mid F)$. Let V be the underlying $\mathbb{Q}$-vector space of B. There exists a non-degenerate alternating $\mathbb{Q}$-bilinear form

$$\Theta : V \times V \to \mathbb{Q}$$

such that $\Theta(bx, y) = \Theta(x, b'y)$ for all $b \in B$, $x, y \in V$. It turns out that G is the group of B-module automorphisms of V preserving the bilinear form. So one has the natural linear representation $\xi_{\mathbb{Q}} : G(\mathbb{Q}) = \mathrm{Aut}_B(V, \Theta) \subset \mathrm{Aut}_{\mathbb{Q}}(V)$.

Note 3: New order correct?

We write $p\mathcal{O}_F = \prod_{i=1}^{r} \mathfrak{p}_i$. By fixing an embedding $\bar{\mathbb{Q}} \to \bar{\mathbb{Q}}_p$, one obtains a bijection between $\Psi = \mathrm{Hom}_{\mathbb{Q}}(F, \bar{\mathbb{Q}})$ and $\coprod_{i=1}^{r} \mathrm{Hom}_{\mathbb{Q}_p}(F_{\mathfrak{p}_i}, \bar{\mathbb{Q}}_p)$. After a rearrangement of indices we can assume that, under the above bijection, τ lies in $\mathrm{Hom}_{\mathbb{Q}_p}(F_{\mathfrak{p}_1}, \bar{\mathbb{Q}}_p)$. We fix the notation $\mathfrak{p} = \mathfrak{p}_1$ for the whole paper. Since p is split in $\mathbb{Q}(\alpha)$ by assumption, $\mathfrak{p}_i\mathcal{O}_K = \mathfrak{q}_i\bar{\mathfrak{q}}_i$ for each i, where the two primes of K over $\mathfrak{p}_i$ are distinguished in such a manner that Ψ_K is bijectively mapped onto $\coprod_{i=1}^{r} \mathrm{Hom}_{\mathbb{Q}_p}(K_{\mathfrak{q}_i}, \overline{\mathbb{Q}}_p)$ under the previous identification map. Now we fix a totally imaginary quadratic extension L of F which is contained in D. Then L splits D globally. Namely there is an isomorphism of L-algebras $D \otimes_F L \cong M_2(L)$. Furthermore one can assume that each $\mathfrak{p}_i$ stays prime in L. So the composite field LK is particularly unramified over p. One writes the prime ideal decomposition as

$$p\mathcal{O}_{LK} = \prod_{i=1}^{r} \mathfrak{q}_i \bar{\mathfrak{q}}_i,$$

and one has a natural isomorphism of $\mathbb{Q}_p$-algebras

$$LK \otimes_{\mathbb{Q}} \mathbb{Q}_p \cong \prod_{i=1}^{r} LK_{\mathfrak{q}_i} \times LK_{\bar{\mathfrak{q}}_i}.$$

Moreover for each i one has an isomorphism $LK_{\mathfrak{q}_i} \otimes_{\mathbb{Q}_p} \mathbb{Q}_p^{ur} \cong \prod_{\mathrm{Hom}_{\mathbb{Q}_p}(LK_{\mathfrak{q}_i}, \overline{\mathbb{Q}}_p)} \mathbb{Q}_p^{ur}$. It is similar for the bar counterpart. Then we obtain an isomorphism of $\mathbb{Q}_p^{ur}$-algebras

$$LK \otimes_{\mathbb{Q}} \mathbb{Q}_p^{ur} \cong \prod_{i=1}^{r} \left(\prod_{\mathrm{Hom}_{\mathbb{Q}_p}(LK_{\mathfrak{q}_i}, \overline{\mathbb{Q}}_p)} \mathbb{Q}_p^{ur} \times \prod_{\mathrm{Hom}_{\mathbb{Q}_p}(LK_{\bar{\mathfrak{q}}_i}, \overline{\mathbb{Q}}_p)} \mathbb{Q}_p^{ur} \right).$$

It induces on the rings of integers a $\mathbb{Z}_p$-algebra isomorphism. One can simplify the notations by using the identification

$$\mathrm{Hom}_{\mathbb{Q}}(LK, \overline{\mathbb{Q}}) = \coprod_{i=1}^{r} \mathrm{Hom}_{\mathbb{Q}_p}(LK_{\mathfrak{q}_i}, \overline{\mathbb{Q}}_p) \times \mathrm{Hom}_{\mathbb{Q}_p}(LK_{\bar{\mathfrak{q}}_i}, \overline{\mathbb{Q}}_p),$$

and the partition $\mathrm{Hom}_{\mathbb{Q}}(LK, \overline{\mathbb{Q}}) = \Phi \coprod \overline{\Phi}$, where $\Phi = \mathrm{Hom}_{\mathbb{Q}}(L, \overline{\mathbb{Q}})$ is identified with the subset of $\mathrm{Hom}_{\mathbb{Q}}(LK, \overline{\mathbb{Q}})$ by extending each embedding of L into $\overline{\mathbb{Q}}$ to an embedding of $LK = L(\alpha) = L + L\alpha$ into $\overline{\mathbb{Q}}$ which is the identity on α. Thus we write the above isomorphism of $\mathbb{Z}_p$-algebras in the form:

$$\prod_{\phi \in \Phi} w(\phi) \times \overline{w}(\phi) : \mathcal{O}_{LK} \otimes_{\mathbb{Z}} W(\mathbb{F}) \xrightarrow{\sim} \prod_{\phi \in \Phi} W(\mathbb{F}) \times W(\mathbb{F}), \tag{1}$$

where for each $\phi \in \Phi$, $w(\phi) \in \mathrm{Hom}_{\mathbb{Q}_p}(LK_{\mathfrak{q}_i}, \overline{\mathbb{Q}}_p)$ and $\overline{w}(\phi) \in \mathrm{Hom}_{\mathbb{Q}_p}(LK_{\bar{\mathfrak{q}}_i}, \overline{\mathbb{Q}}_p)$ for certain i. By abuse of notations we also write the character $w(\phi)$ (resp. $\overline{w}(\phi)$) as ϕ (resp. $\bar{\phi}$) simply. In the following we come to an important notion for this section.

Definition 2.1. Let S be an $\mathcal{O}_{F_\mathfrak{p}}$-scheme and $\mathscr{E}$ be a locally free coherent $\mathcal{O}_S$-sheaf. It is said to be a sheaf of type (L, Ψ_K) if $\mathcal{O}_{LK} \subset \mathrm{End}_{\mathcal{O}_S}(\mathscr{E})$ and $\mathscr{E} \otimes_{\mathbb{Z}} W(\mathbb{F})$ has a decomposition induced by the isomorphism (1) as follows:

$$\mathscr{E} \otimes W(\mathbb{F}) = \bigoplus_{\phi \in \Phi} (\mathscr{E}_\phi \oplus \mathscr{E}_{\bar{\phi}}), \tag{2}$$

where $\mathscr{E}_\phi$ corresponds to the character $w(\phi)$ and $\mathscr{E}_{\bar{\phi}}$ corresponds to the character $\overline{w}(\phi)$ with the rank condition: $\mathscr{E}_\phi$ and $\mathscr{E}_{\bar{\phi}}$ are of rank one if $\phi|_F = \tau$, while $\mathscr{E}_\phi$ is of rank two and $\mathscr{E}_{\bar{\phi}} = 0$ if $\phi|_F \neq \tau$.

In order to define a level structure one shall choose an integral structure of the $\mathbb{Q}$-vector space V. One chooses an order $\mathcal{O}_D$ of D containing $\mathcal{O}_L$ with certain additional properties (see [26], §2), and put $\mathcal{O}_B = \mathcal{O}_D \otimes_{\mathcal{O}_F} \mathcal{O}_K$ (so $\mathcal{O}_{LK} \subset \mathcal{O}_B$). Then one takes the lattice $V_{\mathbb{Z}}$ of V to be the free $\mathbb{Z}$-module $\mathcal{O}_B$ and puts $G(\mathbb{Z}) = \mathrm{Aut}_{\mathcal{O}_B}(V_{\mathbb{Z}}, \Theta)$. Thus one has an integral structure $\xi : G(\mathbb{Z}) \to \mathrm{Aut}_{\mathbb{Z}}(V_{\mathbb{Z}})$ of the $\mathbb{Q}$-algebraic group morphism $\xi_{\mathbb{Q}}$.

Proposition 2.2 (Proposition 2.14 and Corollary 3.14 in [26]). *For every level structure $C = C_p \times C^p \subset G(\mathbb{A}_f)$ with $C_p = G(\mathbb{Z}_p)$ and C^p small enough, there exists a proper $\mathcal{O}_{F_{\mathfrak{p}}}$-scheme $\mathscr{M}_C$ which is the coarse moduli space of certain moduli functor of PEL type* (*see Proposition* 2.14 *in* [26] *for the description of the moduli functor*) *with the endomorphism algebra $\mathcal{O}_B$. Furthermore, if D is assumed to be split at $\mathfrak{p}$, then the reduction $\mathscr{M}_C$ modulo p is smooth over $\mathbb{F}$.*

We take one of the geometrically connected components $\mathscr{M}$ of $\mathscr{M}_C$ with the reduction $\mathscr{M}_0$ modulo p. For our purpose we shall take C^p small enough so that we have the universal abelian scheme $f : \mathscr{X} \to \mathscr{M}$. Under this assumption the genus of $\mathscr{M}$ must be strictly greater than one. By the construction of the moduli functor, the injection $\mathcal{O}_B \hookrightarrow \mathrm{End}_{\mathscr{M}}(\mathscr{X})$ turns $R^1 f_* \mathcal{O}_{\mathscr{X}}$ into a sheaf of (L, Ψ_K)-type.

3. Dieudonné modules and Newton polygons

Let A be an abelian variety which is represented by an $\mathbb{F}$-rational point of $\mathscr{M}_0$. Let $(\mathbb{D} = \mathbb{D}(A), \mathscr{F}, \mathscr{V})$ be the associated (contravariant) Dieudonné module. $\mathbb{D}$ is a free $W(\mathbb{F})$-module of rank $8n$ and one has the identifications of k-vector spaces:

$$\mathbb{D}/p\mathbb{D} = H^1_{dR}(A), \quad \mathscr{V}\mathbb{D}/p\mathbb{D} = H^0(A, \Omega_A), \quad \mathbb{D}/\mathscr{V}\mathbb{D}(A) = H^1(A, \mathcal{O}_A).$$

In this section we shall analyze the structure of $\mathbb{D}$ in the presence of the endomorphism structure. Actually, since $\mathcal{O}_B \subset \mathrm{End}(A)$, it follows that $\mathcal{O}_B \subset \mathrm{End}(\mathbb{D})$ and particularly $\mathcal{O}_{LK} \subset \mathrm{End}(\mathbb{D})$. Therefore $\mathbb{D}$ is an $\mathcal{O}_{LK} \otimes_{\mathbb{Z}} W(\mathbb{F})$-module. The isomorphism (1) in Section 2 gives the decomposition

$$\mathbb{D} = \bigoplus_{\phi \in \Phi} (\mathbb{D}_\phi \oplus \mathbb{D}_{\bar{\phi}}).$$

We put for $1 \leqq i \leqq r$ the local degree $f_i = [F_{\mathfrak{p}_i} : \mathbb{Q}_p]$ and $L_{\mathfrak{p}_i} = L \otimes_F F_{\mathfrak{p}_i}$. For an element $\phi \in \Phi$ one defines $\phi^* \in \Phi$ to be the other element whose restriction to F is the same as that of ϕ. The following proposition contains the basic properties of each direct summand in the above decomposition.

Proposition 3.1. *The Dieudonné module $\mathbb{D}$ has the following properties*:

(i) *$\mathcal{O}_{LK}$ acts on $\mathbb{D}_\phi$* (*resp. $\mathbb{D}_{\bar{\phi}}$*) *via the character $w(\phi)$* (*resp. $\overline{w}(\phi)$*).

(ii) *There is an endomorphism* $\Pi \in \mathcal{O}_B \otimes \mathbb{Z}_p$ *which induces a morphism* $\Pi : \mathbb{D}_\phi \to \mathbb{D}_{\phi^*}$. *It is an isomorphism for* $\phi|_F = \tau$.

(iii) *For* $\phi|_F \neq \tau$, $\mathscr{F}(\mathbb{D}_\phi) = \mathbb{D}_{\sigma\phi}$. *For* $\phi|_F = \tau$, $p\mathbb{D}_{\sigma\phi} \subsetneqq \mathscr{F}\mathbb{D}_\phi \subsetneqq \mathbb{D}_{\sigma\phi}$.

(iv) *For each* $\phi \in \Phi$, *both* $\mathbb{D}_\phi$ *and* $\mathbb{D}_{\bar{\phi}}$ *are of rank* 2.

(v) *The polarization induces a perfect alternative pairing between* $\mathbb{D}_\phi$ *and* $\mathbb{D}_{\bar{\phi}}$. *Moreover* $\mathbb{D}_\phi \perp \mathbb{D}_{\phi'}$ *unless* $\phi' = \bar{\phi}$. *Thus* $\mathbb{D}_\phi$ *and* $\mathbb{D}_{\bar{\phi}}$ *are dual to each other.*

Proof. (i) follows from the definition. The existence of $\Pi \in \mathcal{O}_B \otimes \mathbb{Z}_p$ with the property as in (ii) is actually a part of the conditions on $\mathcal{O}_B$ (see [26], §2). Clearly $\mathscr{F}$ commutes with the $\mathcal{O}_B$-action on $\mathbb{D}$. Since $\mathscr{F}$ is σ-semilinear, one has $\mathscr{F}(\mathbb{D}_\phi) \subseteqq \mathbb{D}_{\sigma\phi}$ by (i). Similarly one has $\mathscr{V}(\mathbb{D}_{\sigma\phi}) \subseteqq \mathbb{D}_\phi$. For a fixed ϕ we consider the short exact sequence

$$0 \to \frac{\mathscr{V}(\mathbb{D}_{\sigma\phi})}{p\mathbb{D}_\phi} \to \frac{\mathbb{D}_\phi}{p\mathbb{D}_\phi} \to \frac{\mathbb{D}_\phi}{\mathscr{V}\mathbb{D}_{\sigma\phi}} \to 0.$$

By the rank condition in Definition 2.1, one has $\dim_{\mathbb{F}} \frac{\mathbb{D}_\phi}{\mathscr{V}\mathbb{D}_{\sigma\phi}}$ is equal to one if $\phi|_F = \tau$, and equal to two if $\phi|_F \neq \tau$. Moreover $\dim_{\mathbb{F}} \frac{\mathbb{D}_{\bar{\sigma}\phi}}{\mathscr{V}\mathbb{D}_{\bar{\phi}}}$ is equal to one in the former case and zero in the latter case. By duality, namely $H^0(A, \Omega_A)^* \cong H^1(A^t, \mathcal{O}_{A^t})$ with A^t the dual abelian variety of A, it follows that $\dim_{\mathbb{F}} \frac{\mathscr{V}(\mathbb{D}_{\sigma\phi})}{p\mathbb{D}_\phi}$ is equal to one in the former case and equal to zero in the latter case. So in both cases $\dim_{\mathbb{F}} \frac{\mathbb{D}_\phi}{p\mathbb{D}_\phi} = 2$ and therefore $\mathrm{rank}_{W(\mathbb{F})} \mathbb{D}_\phi = 2$. (iv) follows from (v). By the above proof, we have $\mathscr{V}(\mathbb{D}_{\sigma\phi}) = p\mathbb{D}_\phi$ for $\phi|_F \neq \tau$ and $p\mathbb{D}_\phi \subsetneqq \mathscr{V}(\mathbb{D}_{\sigma\phi}) \subsetneqq \mathbb{D}_\phi$ otherwise. By applying $\mathscr{F}$ to both sides and dividing by p if necessary, one obtains (iii). Finally, since $\psi(lx, y) = \psi(x, l'y)$ for all $x, y \in \mathbb{D}$ and $l \in B$. We take two idempotents $l_\phi, l_{\phi'}$ with $l_\phi \in (\mathcal{O}_{LK})_\phi$ and $l_{\phi'} \in (\mathcal{O}_{LK})_{\phi'}$. Then $\psi(l_\phi x, l_{\phi'} y) = \psi(x, l'_\phi l_{\phi'} y) = 0$ unless $\bar{\phi} = \phi'$, since $l'_\phi \in (\mathcal{O}_{LK})_{\bar{\phi}}$. Thus (v) follows.

In the following we determine the possible Newton polygons of $\mathbb{D}$. As it is an isogeny invariant, we introduce the (F-)σ-isocrystal $(N = \mathbb{D} \otimes_{\mathbb{Z}} \mathbb{Q}, \mathscr{F})$, and similarly for $\phi \in \Phi$, the direct summand N_ϕ (resp. $N_{\bar{\phi}}$) which itself is generally not a sub F-σ-isocrystal by the previous proposition. For each $1 \leqq i \leqq r$, we put $N_i = \bigoplus_{\phi \in \mathrm{Hom}_{\mathbb{Q}_p}(L_{\mathfrak{p}_i}, \bar{\mathbb{Q}}_p)} N_\phi$ and similarly for $N_{\bar{i}}$. Then N_i and $N_{\bar{i}}$ are indeed F-σ-isocrystals for each i. However if we set $\mathscr{F}_i = (\mathscr{F}^{f_i}\Pi)|_{N_i}$, then by the relation $\sigma^{f_i}\phi = \phi^*$, one sees that $(N_\phi, \mathscr{F}_i)$ is indeed an F-σ^{f_i}-isocrystal for each $\phi \in \mathrm{Hom}_{\mathbb{Q}_p}(L_{\mathfrak{p}_i}, \bar{\mathbb{Q}}_p)$, and similarly for the bar counterpart.

Proposition 3.2. *Let* $(N_i, \mathscr{F})$ *be the F-σ-isocrystal as above. Then it has the following possible Newton slopes*:

(i) *For* $i = 1$ *the Newton slopes are either* $4f_1 \times 1/2f_1$ *or* $2f_1 \times (0, 1/f_1)$.

(ii) *For* $i \geqq 2$ *the Newton slope is* $4f_i \times 0$.

Proof. Since $\mathscr{F} : N_\phi \to N_{\sigma\phi}$ is an isogeny of isocrystals, $(N_\phi, \mathscr{F}_i)$ and $(N_{\sigma\phi}, \mathscr{F}_i)$ have the same Newton slopes. So the computation is reduced to the σ^{f_i}-isocrystals $(N_\phi, \mathscr{F}_i)$ of height 2. Thus the result follows easily from the classification of the isocrystals of height 2 over $\mathbb{F}$ (cf. Lemma 4.4 in [26] or [31]). □

We put $d = [F_{\mathrm{p}} : \mathbb{Q}_p] = f_1$. The following corollary follows easily from the last proposition.

Corollary 3.3. *Let A be an abelian variety which is represented by an* $\mathbb{F}$*-rational point of* $\mathscr{M}_0$. *Then the Newton polygon of A is of the following two possible types*:

$$(4n - 2d) \times 0, \quad 2d \times 1/d, \quad 2d \times (1 - 1/d), \quad (4n - 2d) \times 1;$$

and

$$4(n - d) \times 0, \quad 4d \times 1/2d, \quad 4d \times (1 - 1/2d), \quad 4(n - d) \times 1.$$

Proof. It suffices to notice that N_i and $N_{\bar{i}}$ are dual to each other as σ-isocrystals by Proposition 3.1 (v) for each i. □

Remark 3.4. We see that there are only two possible Newton polygons for closed $\mathbb{F}$-points of the moduli space. The existence of the abelian varieties with the given Newton polygons was shown by Honda–Tate theory. We refer to [1] or [26] for the details.

4. The decomposition of the Higgs bundle over a Shimura curve in char *p*

Let $f : \mathscr{X} \to \mathscr{M}$ be the universal abelian scheme in Section 2. By abuse of notations we denote it again by f the base change to $\overline{\mathbb{Z}}_p$. Let $f^0 : \mathscr{X}^0 \to \mathscr{M}^0$ be the base change of f to $\overline{\mathbb{Q}}_p$ and $f_0 : \mathscr{X}_0 \to \mathscr{M}_0$ be the base change to $\mathbb{F}$. Let $\mathscr{H}^1_{dR} = R^1 f_{0*}(\Omega^{\cdot}_{\mathscr{X}_0 \mid \mathscr{M}_0}, d)$ be the first relative de Rham bundle over $\mathscr{M}_0$. We put the first Hodge bundle $E^{1,0} = f_{0*}\Omega^1_{\mathscr{X}_0 \mid \mathscr{M}_0}$ and the second Hodge bundle $E^{0,1} = R^1 f_{0*}\mathscr{O}_{\mathscr{X}_0}$. By the E_1-degeneration of the Hodge-to-de Rham spectral sequence, one has the short exact sequence

Note 4: Please check the grammar in "we denote it again by f the base change" (maybe delete "it"?)

$$0 \to E^{1,0} \to \mathscr{H}^1_{dR} \to E^{0,1} \to 0.$$

It is well known that $\mathscr{H}^1_{dR}$ is endowed with the Gauss–Manin connection ∇. By taking the grading $(\mathscr{H}^1_{dR}, \nabla)$ with respect to the Hodge filtration, we obtain the Higgs bundle in char p: $(E, \theta) = (E^{1,0} \oplus E^{0,1}, \theta^{1,0} \oplus \theta^{0,1})$ with $\theta^{0,1} = 0$.

By construction, the endomorphism ring of the universal abelian variety $\mathscr{X}_0$ over $\mathscr{M}_0$ contains $\mathscr{O}_B$. Thus each element $b \in \mathscr{O}_B$ induces a morphism $b : \mathscr{X}_0 \to \mathscr{X}_0$ over $\mathscr{M}_0$. Let $\mathscr{O}_{LK} \subset \mathscr{O}_B$ be the maximal abelian subgroup as in Section 2. We have the following decomposition under the $\mathscr{O}_{LK}$-action.

Proposition 4.1. *The first relative de Rham bundle* $(\mathscr{H}^1_{dR}, \nabla)$ *with the Gauss–Manin connection admits a decomposition into the direct sum of rank two subbundles with an integrable connection*

$$(\mathscr{H}^1_{dR}, \nabla) = \bigoplus_{\phi \in \Phi} (\mathscr{H}^1_{dR,\phi}, \nabla_\phi) \oplus (\mathscr{H}^1_{dR,\bar{\phi}}, \nabla_{\bar{\phi}}),$$

such that $\mathcal{O}_{LK}$ acts on $\mathcal{H}^1_{dR,\phi}$ (resp. $\mathcal{H}^1_{dR,\bar{\phi}}$) via the character $\phi \bmod p$ (resp. $\bar{\phi} \bmod p$). It induces the decomposition of the Higgs bundle into the direct sum of rank two Higgs subbundles

$$(E,\theta) = \bigoplus_{\phi\in\Phi} (E_\phi,\theta_\phi) \oplus (E_{\bar{\phi}},\theta_{\bar{\phi}}).$$

Furthermore, by writing

$$E_\phi = E_\phi^{1,0} \oplus E_\phi^{0,1} \quad \textit{and} \quad E_{\bar{\phi}} = E_{\bar{\phi}}^{1,0} \oplus E_{\bar{\phi}}^{0,1},$$

one has $\operatorname{rank} E_\phi^{0,1} = \operatorname{rank} E_{\bar{\phi}}^{0,1} = 1$ *for* $\phi|_F = \tau$; *while for* $\phi|_F \neq \tau$, *one has* $\operatorname{rank} E_\phi^{0,1} = 2$ *and* $\operatorname{rank} E_{\bar{\phi}}^{0,1} = 0$.

Proof. The decomposition of $\mathcal{H}^1_{dR}$ with respect to the $\mathcal{O}_{LK}$-action follows from Proposition 3.1 (i). Because $\mathcal{O}_{LK}$ acts on $\mathcal{X}_0$ as endomorphisms over $\mathcal{M}_0$, it induces an action on the relative de Rham complex as endomorphisms of complexes. Taking the hypercohomology, it induces an action on the Hodge filtration $0 \subset E^{1,0} \subset \mathcal{H}^1_{dR}$. In other words, $E^{1,0}$ is an $\mathcal{O}_{LK}$-invariant subbundle of $\mathcal{H}^1_{dR}$. Thus one has the corresponding decomposition on $E^{1,0}$. $E^{0,1}$ is the quotient bundle $\dfrac{\mathcal{H}^1_{dR}}{E^{1,0}}$. Then for each ϕ (resp. $\bar{\phi}$), one has an injective morphism $\dfrac{\mathcal{H}^1_{dR,\phi}}{E_\phi^{1,0}} \to E^{0,1}$ (resp. for $\bar{\phi}$) induced by $\mathcal{H}^1_{dR} \twoheadrightarrow E^{0,1}$. So one has an isomorphism

$$E^{0,1} \cong \bigoplus_{\phi\in\Phi} \frac{\mathcal{H}^1_{dR,\phi}}{E_\phi^{1,0}} \oplus \frac{\mathcal{H}^1_{dR,\bar{\phi}}}{E_{\bar{\phi}}^{1,0}}.$$

Denote $\dfrac{\mathcal{H}^1_{dR,\phi}}{E_\phi^{1,0}}$ by $E_\phi^{1,0}$ (similarly for $\bar{\phi}$), we obtain the decomposition of $E^{0,1}$. By the short exact sequence

$$0 \to E_\phi^{1,0} \to \mathcal{H}^1_{dR,\phi} \to E_\phi^{0,1} \to 0,$$

the bundle $E_\phi = E_\phi^{1,0} \oplus E_\phi^{0,1}$ has the same rank as $\mathcal{H}^1_{dR,\phi}$, which is two by Proposition 3.1 (iv). It is similar for $\bar{\phi}$. It is clear that the resulting decomposition on $E^{0,1}$ coincides with the induced action of $\mathcal{O}_{LK}$ on $R^1 f_{0*}\mathcal{O}_{\mathcal{X}_0} = E^{0,1}$ by taking the higher direct image. Since the bundle $E^{0,1}$ is the modulo p reduction of $R^1 f_* \mathcal{O}_{\mathcal{X}}$, it is a sheaf of (L,Ψ_K)-type. The assertions about the ranks of $E_\phi^{0,1}$ and $E_{\bar{\phi}}^{0,1}$ follow from the rank condition in Definition 2.1. Finally the $\mathcal{O}_{LK}$-action decomposes the Gauss–Manin connection as well. In fact, in char 0 one can show that the $\mathcal{O}_{LK}$-action on the relative de Rham bundle is flat with respect to (in other words, commutes with) the Gauss–Manin connection because the endomorphism algebra defines the flat Hodge cycles on the relative Betti cohomology. By reduction modulo p, the $\mathcal{O}_{LK}$-action also commutes with ∇. Because $\mathcal{O}_{LK}$ acts on the direct summands via the characters, ∇ preserves each direct summand in the decomposition. The Higgs field θ on E decomposes accordingly. $\square$

Corollary 4.2. *The Hodge-to-de Rham spectral sequence of the relative de Rham bundle $R^1 f_*(\Omega_{\mathcal{X}|\mathcal{M}}, d)$ degenerates at E_1-level. By taking the grading of $\big(R^1 f_*(\Omega_{\mathcal{X}|\mathcal{M}}, d), \nabla\big)$ with respect to the Hodge filtration, one obtains the Higgs bundle $(\tilde{E},\tilde{\theta})$ over $\mathcal{M}$. The $\mathcal{O}_{LK}$-action on the universal abelian scheme $\mathcal{X}$ over $\mathcal{M}$ as endomorphisms induces a decomposition of*

(AutoPDF V7 22/7/11 09:28) WDG Tmath J-2447 CRELLE, () PMU: K(K) 7/7/2011 pp. 1–26 2447_033-6867 (p. 11)

Higgs bundles

$$(\tilde{E}, \tilde{\theta}) = \bigoplus_{\phi \in \Phi} (\tilde{E}_\phi, \tilde{\theta}_\phi) \oplus (\tilde{E}_{\bar{\phi}}, \tilde{\theta}_{\bar{\phi}}).$$

The modulo p reduction of the above decomposition is the one over $\mathcal{M}_0$ given in Proposition 4.1.

Proof. It suffices to show the E_1-degeneration of the Hodge-to-de Rham spectral sequence of f. It is equivalent to show that the natural morphism $f_*\Omega_{\mathcal{X}\,|\,\mathcal{M}} \to R^1 f_*(\Omega_{\mathcal{X}\,|\,\mathcal{M}}, d)$ is injective. By tensoring with $\mathbb{Q}$, the above morphism is injective by the well-known E_1-degeneration of the Hodge-to-de Rham spectral sequence for the first relative de Rham bundle of f^0 (the generic fiber of f). So the kernel of the morphism consists of only p-torsions. By modulo p and the E_1-degeneration of the closed fiber f_0 of f, there are actually no p-torsions. Thus the Hodge-to-de Rham spectral sequence of f degenerates at E_1-level as well. $\square$

Next we proceed to deduce some basic properties of the Higgs subbundles from the above proposition. According to this proposition, the Higgs subbundle (E_ϕ, θ_ϕ) (resp. $(E_{\bar{\phi}}, \theta_{\bar{\phi}})$) has two nontrivial parts, namely, the $(1,0)$-part and the $(0,1)$-part, if and only if $\phi|_F = \tau$ (resp. $\bar{\phi}|_F = \bar{\tau}$). It is clear that there are totally four such direct summands in the decomposition. We consider them first.

Proposition 4.3. *Let ϕ, ϕ^* be two unique elements of Φ whose restriction to F is equal to τ. Then one has an isomorphism of Higgs bundles $(E_\phi, \theta_\phi) \cong (E_{\phi^*}, \theta_{\phi^*})$. One has also an isomorphism for the bar counterpart.*

Proof. In case of $\phi|_F = \tau$ the endomorphism $\Pi \in \mathcal{O}_B \otimes \mathbb{Z}_p$ of $\mathcal{X}_0$ over $\mathcal{M}_0$ induces the endomorphism $\Pi \in \mathrm{End}(\mathcal{H}^1_{dR})$ which is in fact an automorphism. By restricting Π to each closed point in $\mathcal{M}_0$, one knows from Proposition 3.1 (ii) (modulo p) that it induces an isomorphism $\Pi : \mathcal{H}^1_{dR,\phi} \to \mathcal{H}^1_{dR,\phi^*}$. Since Π commutes with the Gauss–Manin connection and the Hodge filtration, it induces an isomorphism of Higgs bundles by taking the grading with respect to the Hodge filtration:

$$\Pi : (E_\phi, \theta_\phi) \cong (E_{\phi^*}, \theta_{\phi^*}). \quad \square$$

The following result asserts that the chern class of the base Shimura curve $\mathcal{M}_0$ is in fact represented by the second Hodge bundle of the Higgs subbundles appearing in the above proposition. This is one of significant features of the above Higgs subbundles.

Proposition 4.4. *Assume that $p \geqq 2g$. Then for $\phi \in \Phi$ with $\phi|_F = \tau$, the Higgs bundle (E_ϕ, θ_ϕ) in char p is of maximal Higgs field. Consequently one has the equality*

$$c_1(E_\phi^{0,1}) = \frac{1}{2} c_1(\mathcal{M}_0).$$

Analogous statements hold for the bar counterpart.

Proof. The Higgs subbundle $(E_\phi = E_\phi^{1,0} \oplus E_\phi^{0,1}, \theta_\phi)$ is the modulo p reduction of the Higgs bundle $(\tilde{E}_\phi, \tilde{\theta}_\phi)$ over $\mathcal{M}$ by Corollary 4.2. For $\phi|_F = \tau$, the Higgs bundle $(E_\phi^0, \theta_\phi^0)$, that is the base change of $(\tilde{E}_\phi, \tilde{\theta}_\phi)$ to $\mathcal{M}^0$, is actually of maximal Higgs field (see [30]). That is, the Higgs field ${\theta_\phi^0}^{1,0} : {E_\phi^0}^{1,0} \to {E_\phi^0}^{0,1} \otimes \Omega_{\mathcal{M}^0}$ is an isomorphism. Then under the assumption on p, we claim that the Higgs field in char p must be maximal. In fact, the Higgs field $\theta_\phi^{1,0}$ can not be zero. Otherwise, the Higgs subbundle $(E_\phi^{1,0}, 0)$ of (E_ϕ, θ_ϕ) is of non-positive degree by the Higgs semistability (see Proposition 6.1). This is in contradiction with the fact that

$$\deg E_\phi^{1,0} = \deg {E_\phi^0}^{1,0} = \frac{1}{2}\deg \Omega_{\mathcal{M}^0} = g - 1 > 0.$$

Since $\theta_\phi^{1,0}$ is a nonzero morphism between line bundles with the same degree, it must be an isomorphism.

Moreover ${E_\phi^0}^{0,1}$ is isomorphic to the dual ${E_\phi^0}^{1,0*}$ of ${E_\phi^0}^{1,0}$. By the theorem of Langton ([19], Main Theorem A′), the isomorphism extends to an isomorphism $E_\phi^{0,1} \cong E_\phi^{1,0*}$. Thus the maximality of θ_ϕ implies that $(E_\phi^{0,1})^2 \cong \mathcal{T}_{\mathcal{M}_0}$. Taking the cycle classes of both sides of the isomorphism, we obtain the claimed formula. □

We believe that one can remove the condition on the prime p in the above proposition. It is clear that these Higgs subbundles in the decomposition of (E, θ) are divided into two types: one is of maximal Higgs field and the other is of zero Higgs field. This is the char p analogue of the corresponding result in [30] in the char 0 case.

5. The Newton polygon jumping locus of the Shimura curve

In this section we prove the mass formula for the Shimura curve $\mathcal{M}_0$. We refer to [7] for the definition of the Newton polygon stratifications and other related notions. By Corollary 3.3, there are only two possible Newton polygons for points in $\mathcal{M}_0(\mathbb{F})$. We denote by $\mathcal{S}$ the subset of $\mathcal{M}_0(\mathbb{F})$ consisting of the closed points for which the Newton polygon jumps. By a theorem of Grothendieck–Katz (see [15]), the Newton polygon jumps under specialization, and $\mathcal{S}$ is an algebraically closed subset of $\mathcal{M}(\mathbb{F})$. In particular, the cardinality of $\mathcal{S}$ is finite.

We find that the morphisms $\mathcal{F}^n_{\mathcal{X}_0 \mid \mathcal{M}_0} : F^{n*}_{\mathcal{M}_0} E^{0,1} \to E^{0,1}$, $n \geqq 1$, where $\mathcal{F}^n_{\mathcal{X}_0 \mid \mathcal{M}_0}$ is the composition of relative Frobenius morphisms (see [15] and [21]), can be applied to compute the number $|\mathcal{S}|$, as is very interesting. One notices that the restriction of $\mathcal{F}_{\mathcal{X}_0 \mid \mathcal{M}_0}$ induces a morphism $\mathcal{F}_{\mathcal{X}_0 \mid \mathcal{M}_0} : F^*_{\mathcal{M}_0} E_\phi^{0,1} \to E_{\sigma\phi}^{0,1}$ for each $\phi \in \Phi$, since the Frobenius morphism on $\mathbb{D}$ is σ-semilinear. Since each prime of F is inert in L, we shall use the same letter $\mathfrak{p}$ to denote the prime of L lying over the prime $\mathfrak{p}$ of F. We write the subset of Φ as

$$\mathrm{Hom}_{\mathbb{Q}_p}(L_\mathfrak{p}, \overline{\mathbb{Q}}_p) = \{\phi_1, \ldots, \phi_d, \phi_1^*, \ldots, \phi_d^*\},$$

in such a way that $\phi_1|_F = \phi_1^*|_F = \tau$, and the Frobenius automorphism σ, which is the generator of $\mathrm{Gal}(L_\mathfrak{p} \mid \mathbb{Q}_p)$, acts on the set as the cyclic permutation of $2d$ letters. For example, $\sigma\phi_d = \phi_1^*$, $\sigma\phi_d^* = \phi_1$, and so on.

Proposition 5.1. *The notations are as above and all morphisms in the following are the relative Frobenius morphisms. Let $\phi \in \Phi$. Then the following statements hold*:

(i) *For* $\phi \in \mathrm{Hom}_{\mathbb{Q}_p}(L_{\mathfrak{p}}, \overline{\mathbb{Q}}_p)$, *one has two possibilities*:

(i.1) *If* $d = 1$, *then* $F^*_{\mathcal{M}_0} E^{0,1}_{\phi_1} \to E^{0,1}_{\phi_1^*}$ *is nonzero. The same holds for* $F^*_{\mathcal{M}_0} E^{0,1}_{\phi_1^*} \to E^{0,1}_{\phi_1}$ *and the bar counterparts. Moreover, they have the same zero locus.*

(i.2) *If* $d \geqq 2$, *then* $F^*_{\mathcal{M}_0} E^{0,1}_{\phi_1} \to E^{0,1}_{\phi_2}$ *is injective*, $F^*_{\mathcal{M}_0} E^{0,1}_{\phi_d} \to E^{0,1}_{\phi_1^*}$ *is surjective, and for the bar counterparts,*

$$(\mathcal{F}_{\mathcal{X}_0 \mid \mathcal{M}_0})|_{F^*_{\mathcal{M}_0} E^{0,1}_{\bar{\phi}_1}} = (\mathcal{F}_{\mathcal{X}_0 \mid \mathcal{M}_0})|_{F^*_{\mathcal{M}_0} E^{0,1}_{\bar{\phi}_1^*}} = 0.$$

Moreover if $d \geqq 3$, *then* $F^*_{\mathcal{M}_0} E^{0,1}_{\phi_i} \to E^{0,1}_{\phi_{i+1}}$ *is an isomorphism for* $2 \leqq i \leqq d-1$.

(ii) *For* $\phi \notin \mathrm{Hom}_{\mathbb{Q}_p}(L_{\mathfrak{p}}, \overline{\mathbb{Q}}_p)$, *one has an isomorphism*

$$\mathcal{F}_{\mathcal{X}_0 \mid \mathcal{M}_0} : F^*_{\mathcal{M}_0} E^{0,1}_{\phi} \cong E^{0,1}_{\sigma\phi}.$$

Proof. We adopt a pointwise argument here. Let $t \in \mathcal{M}_0(\mathbb{F})$ and let A be the fiber of f_0 over t. Let $(\mathbb{D}(A), \mathcal{F}, \mathcal{V})$ be the Dieudonné module of A. We first discuss the $d \geqq 2$ case. Since $\mathcal{F}_{\mathcal{X}_0 \mid \mathcal{M}_0}(F^*_{\mathcal{M}_0} E^{0,1}_{\bar{\phi}_1}) \subset E^{0,1}_{\sigma\bar{\phi}_1}$ and $E^{0,1}_{\sigma\bar{\phi}_1} = 0$ for $(\sigma\bar{\phi}_1)|_F \neq \bar{\tau}$, we have $\mathcal{F}_{\mathcal{X}_0 \mid \mathcal{M}_0}|_{F^*_{\mathcal{M}_0} E^{0,1}_{\bar{\phi}_1}} = 0$ and similarly for the $\bar{\phi}_1^*$-summand. Now we assume $d \geqq 3$ and look at the morphism $F^*_{\mathcal{M}_0} E^{0,1}_{\phi_i} \cong E^{0,1}_{\phi_{i+1}}$ for $2 \leqq i \leqq d-1$. Without loss of generality we discuss this only for $i = 2$. Since $\phi_i|_F \neq \tau$ for $i = 2, 3$, we have $\dim_{\mathbb{F}} \dfrac{\mathbb{D}(A)_{\phi_i}}{p\mathbb{D}(A)_{\phi_i}} = 2$. Furthermore $\mathcal{V}(\mathbb{D}(A)_{\phi_3}) = p\mathbb{D}(A)_{\phi_2}$ by Proposition 3.1 (iii). We consider the σ-semilinear map

$$\mathcal{F} \bmod p : \frac{\mathbb{D}(A)_{\phi_2}}{\mathcal{V}\mathbb{D}(A)_{\phi_2}} \to \frac{\mathbb{D}(A)_{\phi_3}}{\mathcal{V}\mathbb{D}(A)_{\phi_3}}$$

induced by $\mathcal{F} : \mathbb{D}(A)_{\phi_2} \to \mathbb{D}(A)_{\phi_3}$. It is known that the above map is simply the Hasse–Witt map after identifying the spaces $\dfrac{\mathbb{D}(A)_{\phi_i}}{p\mathbb{D}(A)_{\phi_i}} = H^1(A, \mathcal{O}_A)_{\phi_i}$. Now for $e \in \mathbb{D}(A)_{\phi_2}$, we have that $e \bmod p \in \ker(\mathcal{F} \bmod p)$ if and only if $\mathcal{F}(e) \in p\mathbb{D}(A)_{\phi_3}$, and if and only if $e \in \mathcal{V}(\mathbb{D}(A)_{\phi_3}) = p\mathbb{D}(A)_{\phi_2}$ (by applying $\mathcal{V}$ or $\mathcal{F}$ to both sides). So one sees that $\mathcal{F} \bmod p$ is injective. Hence it must be surjective for the dimensional reason. This proves the isomorphism in (i.2) for $i = 2$ and similarly we have the isomorphisms in (ii). The similar argument proves the injectivity of $F^*_{\mathcal{M}_0} E^{0,1}_{\phi_1} \to E^{0,1}_{\phi_2}$ and the surjectivity of $F^*_{\mathcal{M}_0} E^{0,1}_{\phi_d} \to E^{0,1}_{\phi_1^*}$ follows by duality.

It remains to show the $d = 1$ case. In this case, the result in Corollary 3.3 tells us that the p-rank of abelian varieties in $\mathcal{M}_0(\mathbb{F})$ is either $4(n-1)$ or $4n$. Now by (ii), all of the ϕ-summands with $\phi \notin \mathrm{Hom}_{\mathbb{Q}_p}(L_{\mathfrak{p}}, \overline{\mathbb{Q}}_p)$ have contributions to the p-rank. The existence of closed points of p-rank $4(n-1)$ implies that each of the four morphisms in (i) can not be zero. Moreover, because of the existence of the other p-rank, all of the four morphisms will vanish at a point t as soon as one of them vanishes at t. □

Now we consider the composition of the relative Frobenius morphisms

$$\mathcal{F}^d_{\mathcal{X}_0 \mid \mathcal{M}_0} : E^{0,1\,(p^d)}_{\phi_1} \to E^{0,1\,(p^{d-1})}_{\phi_2} \to \cdots \to E^{0,1\,(p)}_{\phi_d} \to E^{0,1}_{\phi_1^*}.$$

As a consequence of the above analysis of the relative Frobenius morphisms, we have the following

Proposition 5.2. *Let $t \in \mathscr{M}_0(\mathbb{F})$. Then $t \in \mathscr{S}$ if and only if the map*

$$\mathscr{F}^d_{\mathscr{X}_0 \mid \mathscr{M}_0} : E^{0,1\,(p^d)}_{\phi_1} \to E^{0,1}_{\phi_1^*}$$

vanishes at t.

Proof. For simplicity we put $h = \mathscr{F}^d_{\mathscr{X}_0 \mid \mathscr{M}_0}$. The $d = 1$ case follows directly from Proposition 5.1 (i). Assume $d \geqq 2$. One direction is clear. Namely if $t \in S$, then the restriction of h to the ϕ_1-summand must be zero at t by Proposition 5.1 (ii). We proceed to show the converse direction. Again by Proposition 5.1 (ii) the ϕ-summands with $\phi \notin \mathrm{Hom}_{\mathbb{Q}_p}(L_{\mathfrak{p}}, \overline{\mathbb{Q}}_p)$ have no contribution to the jumping of the Newton polygons and we only need to consider $\phi \in \mathrm{Hom}_{\mathbb{Q}_p}(L_{\mathfrak{p}}, \overline{\mathbb{Q}}_p)$. It is clear that h maps the ϕ_i-summand to the ϕ_i^*-summand and the ϕ_i^*-summand to the ϕ_i-summand for $1 \leqq i \leqq d$. For example, one has $h(E^{0,1\,(p^d)}_{\phi_1^*}) \subset E^{0,1}_{\phi_1}$. Note that for the ϕ-summand with $\phi|_F \neq \tau$ the rank of h must be reduced by at least one and thus it contributes to the p-rank at most by one. There are $2(d-1)$ such summands in all. By the assumption h vanishes at t on the ϕ_1-summand. Applying the endomorphism Π one sees that h vanishes at t on the ϕ_1^*-summand too. It implies that the p-rank at t is at most $4n - 2d - 2$ by Proposition 5.1 (i.2), and therefore the p-rank can not be $4n - 2d$. Since there are only two possibilities for the p-ranks at $t \in S$. From the above proof, we also see that the restriction of h to each ϕ-summand with $\phi \in \mathrm{Hom}_{\mathbb{Q}_p}(L_{\mathfrak{p}}, \overline{\mathbb{Q}}_p)$ vanishes at t if and only if $t \in S$. □

Note 5: "Since …" is not a complete sentence.

Furthermore, we have

Proposition 5.3. *The zero locus of $\mathscr{F}^d_{\mathscr{X}_0 \mid \mathscr{M}_0} : E^{0,1\,(p^d)}_{\phi_1} \to E^{0,1}_{\phi_1^*}$ is reduced. In other words, the zero divisor of the global section of the line bundle*

$$(E^{0,1\,(p^d)}_{\phi_1})^{-1} \otimes E^{0,1}_{\phi_1^*}$$

defined by $\mathscr{F}^d_{\mathscr{X}_0 \mid \mathscr{M}_0}$ is of multiplicity one.

Before we show this result, we make a digression into the display theory of Dieudonné modules. By a theorem of Serre–Tate ([16]), the equi-characteristic deformation of an abelian variety A in positive characteristic is the same as that of its p-divisible group $A(p)$. The latter is determined by the display of the Dieudonné module $\mathbb{D}\big(A^t(p)\big)$ (see [24] and the references therein). This is also true when polarizations and endomorphisms are considered. We put $\mathbb{D} = \mathbb{D}\big(A^t(p)\big)$ and recall that we have the following decomposition:

$$\mathbb{D} = \bigoplus_{\phi \in \Phi} (\mathbb{D}_\phi \oplus \mathbb{D}_{\bar{\phi}}),$$

where $\Phi = \mathrm{Hom}_{\mathbb{Q}}(L, \overline{\mathbb{Q}})$.

Lemma 5.4. *There exists a basis $\{X_\phi, Y_\phi, X_{\bar{\phi}}, Y_{\bar{\phi}} \mid \phi \in \Phi\}$ of $\mathbb{D}$, such that*

(i) *$\{X_\phi, Y_\phi\}$ is a basis for $\mathbb{D}_\phi$, and $X_{\bar{\phi}}, Y_{\bar{\phi}}$ is the dual basis of $\mathbb{D}_{\bar{\phi}}$.*

(ii) *For $\phi|_F = \tau$, $Y_\phi, X_{\bar{\phi}} \in \mathscr{V}(\mathbb{D})$; while for $\phi|_F \neq \tau$, $X_{\bar{\phi}}, Y_{\bar{\phi}} \in \mathscr{V}(\mathbb{D})$.*

(AutoPDF V7 22/7/11 09:28) WDG Tmath J-2447 CRELLE, () PMU: K(K) 7/7/2011 pp. 1–26 2447_033-6867 (p. 15)

Proof. Note that for $\phi \in \Phi$, we have a short exact sequence:

$$0 \to \mathscr{V}\mathbb{D}_{\sigma\phi}/p\mathbb{D}_{\phi} \to \mathbb{D}_{\phi}/p\mathbb{D}_{\phi} \to \mathbb{D}_{\phi}/\mathscr{V}\mathbb{D}_{\sigma\phi} \to 0.$$

For $\phi|_F \neq \tau$, the statements in (i) and (ii) are obvious, by Proposition 3.1. For $\phi|_F = \tau$, first of all, we prove that we can choose a basis X_ϕ, Y_ϕ for $\mathbb{D}_\phi$, such that $Y_\phi \in \mathscr{V}(\mathbb{D}_{\sigma\phi})$. In this case, by Proposition 3.1, the dimensions of the terms appearing in the above exact sequence are in turn 1, 2, 1. Let $X_\phi, Y_\phi \in \mathbb{D}_\phi$, such that the image of $\overline{X}_\phi$ generates $\mathbb{D}_\phi/\mathscr{V}\mathbb{D}_{\sigma\phi}$ as vector space, and $\overline{Y}_\phi$ generates $\mathscr{V}(\mathbb{D}_{\sigma\phi})$ as vector space. Then $\overline{X}_\phi$, $\overline{Y}_\phi$ generate $\mathbb{D}_\phi/p\mathbb{D}_\phi$. By Nakayama's lemma, X_ϕ, Y_ϕ is a basis of $\mathbb{D}_\phi$ with $Y_\phi \in \mathscr{V}(\mathbb{D}_{\sigma\phi})$. Secondly, we prove that there is a dual basis $X_{\bar\phi}$, $Y_{\bar\phi}$ of $\mathbb{D}_{\bar\phi}$, such that $X_{\bar\phi} \in \mathscr{V}(\mathbb{D})$. Similarly as above, we can find a basis x, y of $\mathbb{D}_{\bar\phi}$ with $y \in \mathscr{V}(\mathbb{D})$. Let

$$H = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

be the intersection matrix of X_ϕ, Y_ϕ and x, y. Thus we have the valuation $v_p(d) > 0$, b, c are invertible and H is invertible. By solving a system of linear equations, we see that $X_{\bar\phi} = \dfrac{1}{\det(H)}(dx - cy)$ and $Y_{\bar\phi} = \dfrac{1}{\det(H)}(-bx + ay)$ satisfy the requirements, since $v_p(d) > 0$ implies that $dx = p\,d'x = \mathscr{V}\mathscr{F}(d'x) \in \mathscr{V}(\mathbb{D})$. □

Under this basis of $\mathbb{D}$, the corresponding display is

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

where the matrix A, C are (we take $n = 2$, $d = [F_{\mathfrak{p}} : \mathbb{Q}_p] = 2$ for example)

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & a_1 & c_1 & 0 & 0 \\ a_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ b_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & a_{1^*} & c_{1^*} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_{2^*} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & b_{2^*} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

and

$$C = \begin{pmatrix} 0 & 0 & 0 & 0 & b_1 & d_1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & b_{1^*} & d_{1^*} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Here the basis is arranged in an obvious manner. In this case, the Frobenius is given by the matrix $\begin{pmatrix} A & pB \\ C & pD \end{pmatrix}$.

Lemma 5.5. *Let $R = k[\![t]\!]$. The $\mathcal{O}_B$-action can be extended to* $\operatorname{spec}(R)$. *Hence the display of the infinitesimal universal deformation is given by*

$$\begin{pmatrix} A + TC & p(B + TD) \\ C & pD \end{pmatrix},$$

where T is the Teichmüller lifting of t (that is, $T = (t, 0, \ldots)$). In particular, the matrix $A + TC$, read mod p, is the Hasse–Witt matrix of the deformation corresponding to T.

Proof. It is known from Proposition 2.2 that the local deformation ring of the Shimura curve is regular on one parameter. Let $\mathbb{D}_R$ be the display over R. Then

$$\mathbb{D}_R = \bigoplus_{\phi \in \Phi} (\mathbb{D}_{R,\phi} \oplus \mathbb{D}_{R,\bar{\phi}}),$$

where $\mathbb{D}_{R,\phi}$ (resp. $\mathbb{D}_{R,\bar{\phi}}$) is obtained from $\mathbb{D}_\phi$ (resp. $\mathbb{D}_{\bar{\phi}}$) by extending scalars to $W(R)$, with the naturally given action of $W(k)$ on each component. Recall that the action of $\mathcal{O}_{LK}$ is defined via the map

$$\mathcal{O}_{LK} \to \bigoplus_{\phi} \big(W(k) \oplus W(k)\big), \quad a \mapsto \big(\ldots, w(\phi)(a), \overline{w}(\phi)(a), \ldots\big).$$

This is a map of Dieudonné modules if and only if it commutes with the Frobenius; that is, if and only if

$$M_1 M_2^\sigma = M_2 M_1,$$

where

$$M_1 = \begin{pmatrix} A + TC & p(B + TD) \\ C & pD \end{pmatrix},$$

$$M_2 = \begin{pmatrix} \operatorname{diag}\big(\ldots, w(\phi)(a), \ldots, \overline{w}(\phi)(a), \ldots\big) & 0 \\ 0 & \operatorname{diag}\big(\ldots, w(\phi)(a), \ldots, \overline{w}(\phi)(a), \ldots\big) \end{pmatrix}.$$

It is easy to verify that this is true, by a direct computation. □

Now we come to the proof of Proposition 5.3.

Proof. The universal Dieudonné module $\mathbb{D}_R$ is displayed by the matrix

$$A = \begin{pmatrix} A + TC & B + TD \\ C & D \end{pmatrix}.$$

Now we show that the locus of $\mathscr{F}^d$ is of multiplicity one. For this we put

$$\mathscr{V}^d = \mathscr{V}^d_{\mathscr{X}_0 \mid \mathscr{M}_0} : E^{1,0}_{\bar{\phi}_1^*} \to E^{0,1\,p^d}_{\bar{\phi}_1},$$

which is the dual of $\mathscr{F}^d$. It is then equivalent to show that the zero locus of $\mathscr{V}^d$ is of multiplicity one. Moreover, for simplicity we just take $n = d = 2$ in the following argument, and the argument for the general case is completely the same. The following matrix mod p is the Hasse–Witt matrix of the Frobenius of the local deformation:

$$A + TC = \begin{pmatrix} 0 & 0 & 0 & 0 & a_1 + tb_1 & c_1 + td_1 & 0 & 0 \\ a_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ b_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & a_{1^*} + tb_{1^*} & c_{1^*} + td_{1^*} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_{2^*} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & b_{2^*} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Thus the Hasse–Witt matrix of $\mathscr{V}^2$ is the following matrix mod p:

$$(A + TC)(A + TC)^\sigma = \begin{pmatrix} 0 & 0 & 0 & f_{14} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & f_{25} & f_{26} & 0 & 0 \\ 0 & 0 & 0 & 0 & f_{35} & f_{36} & 0 & 0 \\ f_{41} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & f_{52} & f_{53} & 0 & 0 & 0 & 0 & 0 \\ 0 & f_{62} & f_{63} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

where

$$\begin{aligned} f_{14} &= (a_1 + tb_1)a_{2^*}^\sigma + (c_1 + td_1)b_{2^*}^\sigma, \\ f_{25} &= a_2(a_1^\sigma + t^\sigma b_1^\sigma), \quad f_{26} = a_2(c_1^\sigma + t^\sigma d_1^\sigma), \\ f_{35} &= b_2(a_1^\sigma + t^\sigma b_1^\sigma), \quad f_{36} = b_2(a_1^\sigma + t^\sigma b_1^\sigma), \\ f_{41} &= (a_{1^*} + tb_{1^*})a_2^\sigma + (c_{1^*} + td_{1^*})b_2^\sigma, \\ f_{52} &= a_{2^*}(a_{1^*}^\sigma + t^\sigma b_{1^*}^\sigma), \quad f_{53} = a_{2^*}(c_{1^*}^\sigma + t^\sigma d_{1^*}^\sigma), \\ f_{62} &= b_{2^*}(a_{1^*}^\sigma + t^\sigma b_{1^*}^\sigma), \\ f_{63} &= b_{2^*}(c_{1^*}^\sigma + t^\sigma d_{1^*}^\sigma). \end{aligned}$$

Thus $\mathscr{V}^2$ is locally given by the function

$$(a_{1^*} + tb_{1^*})a_2^\sigma + (c_{1^*} + td_{1^*})b_2^\sigma = (a_{1^*}a_2^\sigma + c_{1^*}b_2^\sigma) + t(b_{1^*}a_2^\sigma + d_{1^*}b_2^\sigma) = 0.$$

As $t = 0 \in \mathscr{S}$, we have $a_{1^*}a_2^\sigma + c_{1^*}b_2^\sigma = 0$ and $b_{1^*}a_2^\sigma + d_{1^*}b_2^\sigma \neq 0$. Thus the locus is of multiplicity one. □

Theorem 5.6. *Let $\mathscr{M}_0$ be the moduli space constructed in Section 2 and let $\mathscr{S}$ be the Newton polygon jumping locus in $\mathscr{M}_0$. Then in the Chow ring of $\mathscr{M}_0$ the following formula* *holds*:

$$\mathscr{S} = \frac{1}{2}(1 - p^d)c_1(\mathscr{M}_0),$$

where $d = [F_{\mathfrak{p}} : \mathbb{Q}_p]$ *is the local degree. As a consequence, one obtains the mass formula for* $\mathscr{M}_0$:

$$|\mathscr{S}| = (p^d - 1)(g - 1),$$

where g *is the genus of the Shimura curve* $\mathscr{M}_0$.

Proof. The mass formula follows by taking the degree in the cycle formula. By Proposition 4.3 and 4.4, the cycle of the zero locus of $F^d_{\mathscr{X}_0 \mid \mathscr{M}_0} : E^{0,1\,(p^d)}_{\phi_1} \to E^{0,1}_{\phi_1^*}$ is equal to

$$c_1(E^{0,1}_{\phi_1^*}) - c_1(E^{0,1\,(p^d)}_{\phi_1}) = (1 - p^d)c_1(E^{0,1}_{\phi_1}) = \frac{1}{2}(1 - p^d)c_1(\mathscr{M}_0).$$

Then the theorem follows from Proposition 5.2 and 5.3. □

Remark 5.7. The last two sections have certain overlaps with parts of the paper [12] by P. Kassaei. In particular one shall compare Corollary 4.4 and Proposition 5.3 here with Proposition 4.1 and 4.3 in [12].

6. Stability and instability of the Higgs subbundles in char *p*

In this section we study the stability and instability of the Higgs subbundles constructed in Proposition 4.1. We will assume that $p \geqq 2g$ in this section, unless otherwise specified.

Proposition 6.1. *With the assumption on* p *as above, we have that for each* $\phi \in \Phi$, (E_ϕ, θ_ϕ) *and* $(E_{\bar\phi}, \theta_{\bar\phi})$ *are Higgs semistable of degree* 0. *Particularly for* $\phi|_F \neq \tau$, *the rank two vector bundles* $E^{0,1}_\phi$ *and* $E^{1,0}_{\bar\phi}$ *are semistable.*

Proof. By construction, for each $\phi \in \Phi$, (E_ϕ, θ_ϕ) and $(E_{\bar\phi}, \theta_{\bar\phi})$ are the modulo p reductions of Higgs bundles in characteristic 0 by Corollary 4.2. By Theorem 4.14 (3) and Proposition 4.19 in [25], they are Higgs semistable under the assumption on p as in the statement. Moreover for each place ϕ with $\phi|_F \neq \tau$ one has $(E_\phi, \theta_\phi) = (E^{0,1}_\phi, \theta^{0,1}_\phi)$ and $(E_{\bar\phi}, \theta_{\bar\phi}) = (E^{1,0}_{\bar\phi}, \theta^{1,0}_{\bar\phi})$ by Proposition 4.1. The Higgs field $\theta^{0,1}_\phi$ is by definition zero, and $\theta^{1,0}_{\bar\phi}$ is also zero as $E^{0,1}_{\bar\phi}$ is a zero bundle. □

For a semistable bundle E of rank r over a smooth projective curve C defined over $\mathbb{F}$, one can ask further the semistability of the bundle over C under the n-th iterated Frobenius pull-back $F^{*n}_C E$ for $n \geqq 1$. It turns out that the bundles $F^{*n}_C E$ are not necessarily semistable. In order to measure the instability of $F^*_C E$ one introduces and studies the invariant $v(F^*_C E) = \mu_{\max}(F^*_C E) - \mu_{\min}(F^*_C E)$ where $\mu_{\max}(F^*_C E)$ (resp. $\mu_{\min}(F^*_C E)$) is the slope of E_1

Note 6:
Do you mean "one can further ask about the semistability"? (e.g. insert "about")

(resp. $\dfrac{E_n}{E_{n-1}}$) in the Harder–Narasimhan filtration of F_C^*E:

$$0 = E_0 \subset E_1 \subset \cdots \subset E_n = F_C^*E.$$

The following result says that F_C^*E can not be very instable by exhibiting an upper bound of $\nu(F_C^*E)$.

Theorem 6.2 (Lange–Stuhler [17], Satz 2.4 for $r = 2$; Shepherd-Barron [27], Corollary 2, and Sun [29], Theorem 3.1 for arbitrary rank). *Let E be a rank r semistable bundle over a smooth projective curve C of genus g defined over $\mathbb{F}$. Then one has the inequality*

$$\nu(F_C^*E) \leqq (r-1)(2g-2).$$

In particular, F_C^*E is still semistable when $g \leqq 1$.

Based on the above inequality one can also deduce a generalization of it for $n \geqq 2$ (see [21], Theorem 3.7). A. Langer ([18], Corollary 6.2) actually obtained a better bound on this issue and generalized the above inequality as well to a higher dimensional base. It is then interesting to find examples where the upper bound of the inequality is reached. In the case where the local degree $[F_{\mathfrak{p}} : \mathbb{Q}_p]$ is strictly larger than one, certain Higgs subbundles over the Shimura curve $\mathscr{M}_0$ do provide such examples (see the proposition in [11], §4.4, for a classification of semistable bundles of rank two over curves in char 2).

Proposition 6.3. *Assume that $[F_{\mathfrak{p}} : \mathbb{Q}_p] > 1$. The rank two semistable bundles $E^{0,1}_{\phi_d}$ and $E^{0,1}_{\phi_d^*}$ over $\mathscr{M}_0$ achieve the upper bound in Theorem 6.2. That is, one has the equality*

$$\nu(E^{0,1\,(p)}_{\phi_d}) = \nu(E^{0,1\,(p)}_{\phi_d^*}) = 2g-2.$$

Proof. It suffices to show the result for the ϕ_d piece. The proof for another piece is completely similar. We consider the morphism

$$F_{\mathscr{X}_0\,|\,\mathscr{M}_0} : E^{0,1\,(p)}_{\phi_d} \to E^{0,1}_{\phi_1^*},$$

where $E^{0,1}_{\phi_1^*}$ is a line bundle since $\phi_1^*|_F = \tau$. We put then $\mathscr{E}$ (resp. $\mathscr{E}'$) to be the kernel (resp. the image) of the above morphism. That is, we have the following short exact sequence:

$$0 \to \mathscr{E} \to E^{0,1\,(p)}_{\phi_d} \to \mathscr{E}' \to 0.$$

Since $\deg(E^{0,1}_{\phi_d}) = 0$ and $\deg(E^{0,1}_{\phi_1^*}) = 1-g$, we have $\deg E^{0,1\,(p)}_{\phi_d} = 0$ and $\deg \mathscr{E}' \leqq 1-g$. So by the inequality in Theorem 6.2, one has the following inequalities:

$$\mu(\mathscr{E}) - \mu(\mathscr{E}') \leqq \nu(E^{0,1\,(p)}_{\phi_d}) \leqq 2g-2 \leqq \mu(\mathscr{E}) - \mu(\mathscr{E}').$$

It follows that the above inequalities have to be an equality at each step and particularly the assertion of the proposition follows. $\square$

Combining this proposition with Proposition 5.1 (ii), one obtains Theorem 1.4 (iii).

The other extreme about the semistability of a vector bundle under iterated Frobenius pull-backs is expressed in the following definition.

Definition 6.4. Let E be a semistable vector bundle over a smooth projective curve C as above. It is said to be strongly semistable if for all $n \geqq 1$, the $F_C^{*n}E$ are semistable.

In the case that for certain $n \geqq 1$ one has an isomorphism $F_C^{*n}E \cong E$, the bundle E is obviously strongly semistable. The following theorem gives us a characterization of the category of the strongly semistable bundles over C.

Theorem 6.5 (Lange–Stuhler [17], §1). *Let C be a smooth projective curve as above and $\pi_1^{et}(C)$ be its étale fundamental group. Let E be a vector bundle over C of rank r. Then the following conditions about E are equivalent*:

(i) *There exists $n \geqq 1$ such that $F_C^{*n}E \cong E$.*

(ii) *There exists an étale covering map $\pi : \tilde{C} \to C$ such that π^*E over $\tilde{C}$ is trivial.*

(iii) *E corresponds to a continuous representation $\pi_1^{et}(C) \to Gl_r(\mathbb{F})$ where $Gl_r(\mathbb{F})$ is equipped with the discrete topology.*

One calls the bundle E satisfying one of the above equivalent conditions étale trivializable. The bundle E is strongly semistable if and only if there exists $n \geqq 0$ such that $F_C^{*n}E$ is étale trivializable.

We can find such examples again among the Higgs subbundles in this study.

Proposition 6.6. (i) *Assume p is not inert in F. Then for $\phi \notin \mathrm{Hom}_{\mathbb{Q}_p}(L_{\mathfrak{p}}, \overline{\mathbb{Q}}_p)$, $E_{\phi}^{0,1}$ and $E_{\bar{\phi}}^{1,0}$ are étale trivializable, and particularly strongly semistable.* Note 7: Insert "the bundles" before "$E_{\phi}^{0,1}$ and $E_{\bar{\phi}}^{1,0}$"?

(ii) *In case $d = [F_{\mathfrak{p}} : \mathbb{Q}_p] > 1$, the semistable bundles $E_{\phi_i}^{0,1}$ for $2 \leqq i \leqq d$ are not strongly semistable and consequently stable.*

Proof. (i) By Proposition 5.1 (ii), the morphism $\mathscr{F}_{\mathscr{X}_0 \mid \mathscr{M}_0} : E_{\phi}^{0,1\,(p)} \to E_{\sigma\phi}^{0,1}$ is an isomorphism in the case that $\phi|_F \neq \tau$. When $\phi \notin \mathrm{Hom}_{\mathbb{Q}_p}(L_{\mathfrak{p}_1}, \overline{\mathbb{Q}}_p)$, we have $(\sigma\phi)|_F \neq \tau$. This implies that for n large enough the composition of the relative Frobenius morphisms induces an isomorphism $E_{\phi}^{0,1\,(p^n)} \cong E_{\phi}^{0,1}$. The proof for $E_{\bar{\phi}}^{1,0}$ is similar by replacing the relative Frobenius morphism in the argument by the relative Verschiebung morphism $V_{\mathscr{X}_0 \mid \mathscr{M}_0}$.

(ii) For $\phi_i \in \mathrm{Hom}_{\mathbb{Q}_p}(L_{\mathfrak{p}_1}, \overline{\mathbb{Q}}_p)$ with $i \neq 1$, we consider the composition η of the morphisms

$$E_{\phi_i}^{0,1\,(p^{d-i+1})} \overset{F_{\mathscr{X}_0 \mid \mathscr{M}_0}}{\cong} E_{\phi_{i+1}}^{0,1\,(p^{d-i})} \cdots \overset{F_{\mathscr{X}_0 \mid \mathscr{M}_0}}{\cong} E_{\phi_d}^{0,1\,(p)} \twoheadrightarrow E_{\phi_1^*}^{0,1},$$

where the last morphism is surjective by the proof of Proposition 6.3. Thus the kernel of η provides a sub line bundle of positive degree (actually it is equal to $g-1$) in $E_{\phi_i}^{0,1\,(p^{d-i+1})}$ and therefore it is not semistable. The assertion about the stability follows from the following simple lemma. $\square$

Lemma 6.7. *A strictly semistable rank two vector bundle E of degree zero over a smooth projective curve C over $\mathbb{F}$ is strongly semistable.*

Proof. By the assumption, E is an extension of two degree zero line bundles. We write

$$0 \to \mathscr{L}_1 \to E \to \mathscr{L}_2 \to 0.$$

Suppose that $E^{(p)}$ is not semistable, and let $\mathscr{L}$ be a positive sub line bundle of it. Then we have the morphisms

$$\mathscr{L} \to E^{(p)} \to \mathscr{L}_2^{(p)}.$$

Because $\deg \mathscr{L}_2^{(p)} = 0 < \deg \mathscr{L}$, the above composition is zero. Hence the inclusion of $\mathscr{L}$ in $E^{(p)}$ factors through $\mathscr{L}_1^{(p)}$. That is, one has $\mathscr{L} \subset \mathscr{L}_1^{(p)}$. Again because

$$\deg \mathscr{L}_1^{(p)} = 0 < \deg \mathscr{L},$$

we obtain a contradiction. In conclusion, $E^{(p)}$ is semistable. By induction on the iterations of Frobenius pull-backs we see that E is actually strongly semistable. □

7. The Simpson–Ogus–Vologodsky correspondence of the Higgs subbundles in char p

One of the main results in [25] is to establish a char p analogue of the Simpson correspondence. Let $\pi : \mathscr{M}_0' \to \mathscr{M}_0$ be the projection map and $\mathscr{F}_{\mathscr{M}_0} : \mathscr{M}_0 \to \mathscr{M}_0'$ be the relative Frobenius in the commutative diagram of Frobenius morphisms for $\mathscr{M}_0$ over $\mathbb{F}$ (see notations and conventions in Section 1). For the Shimura curve $\mathscr{M}_0$ over $\mathbb{F}$, which is the reduction of a Shimura curve over mixed characteristic, the Cartier transform $C_{\mathscr{M}_0}$ (see [25]) is a functor from the category $MIC(\mathscr{M}_0)$ of flat bundles over $\mathscr{M}_0$ to the category of Higgs bundles $HIG(\mathscr{M}_0')$ over $\mathscr{M}_0'$, with each of them subject to suitable nilpotence conditions. The functor is an equivalence of categories with the quasi-inverse functor $C_{\mathscr{M}_0}^{-1} : HIG(\mathscr{M}_0') \to MIC(\mathscr{M}_0)$. For the full subcategory of flat bundles with vanishing p-curvatures, the functor is just the classical Cartier descent (see [13], §5), and it maps onto the full subcategory of Higgs bundles with trivial Higgs fields over $\mathscr{M}_0'$. It is also clear that the functor transforms the relative de Rham bundle $(\mathscr{H}_{dR}^1, \nabla)$ of the universal family $f_0 : \mathscr{X}_0 \to \mathscr{M}_0$ to $(E', \theta') = \pi^*(E, \theta)$, which is the associated Higgs bundle of the family $f_0' : \mathscr{X}_0' \to \mathscr{M}_0'$, where f_0' is the base change of f_0 via π. In this section we examine the Simpson–Ogus–Vologodsky correspondence and the Cartier transform for the Higgs subbundles in Proposition 4.1. It will be assumed in this section that $p \geqq \max\{2g, 2(n+1)\}$, in order to fulfill the basic requirements in [25], Theorem 3.8. We use $\mathscr{P}'$ to denote the pull-back to $\mathscr{M}_0'$ via π of an algebra-geometric object $\mathscr{P}$ defined over $\mathscr{M}_0$.

We have the ring homomorphisms $\mathcal{O}_{LK} \xrightarrow{\varsigma} \mathrm{End}_{\mathscr{M}_0}(\mathscr{X}_0) \xrightarrow{\prime} \mathrm{End}_{\mathscr{M}_0'}(\mathscr{X}_0')$. Thus ς' denotes the composition morphism. By abuse of notations, the induced ring homomorphisms from $\mathcal{O}_{LK}$ to the endomorphism rings $\mathrm{End}(\mathscr{H}_{dR}^1, \nabla)$ and $\mathrm{End}(E, \theta)$ respectively are again denoted by ς. For example, the $\mathcal{O}_{LK}$-action on (E', θ') via ς' decomposes it into the direct sum of eigen Higgs subbundles

$$(E', \theta') = \bigoplus_{\phi \in \Phi} (E_\phi', \theta_\phi') \oplus (E_{\bar\phi}', \theta_{\bar\phi}'),$$

where $\mathcal{O}_{LK}$ acts on the direct summand E'_ϕ via the character ϕ and similarly for its bar counterpart.

Proposition 7.1. *Let $C_{\mathcal{M}_0} : MIC(\mathcal{M}_0) \to HIG(\mathcal{M}'_0)$ be the Cartier transform. Then for each $\phi \in \Phi$, one has $C_{\mathcal{M}_0}(\mathcal{H}^1_{dR,\phi}, \nabla_\phi) = (E'_\phi, \theta'_\phi)$. The same is true for its bar counterpart.*

Proof. Take $\lambda \in \mathcal{O}_{LK}$. Applying Theorem 3.8 in [25] to the morphism $\varsigma(\lambda) : \mathcal{X}_0 \to \mathcal{X}_0$ and the objects $(\mathcal{O}_{\mathcal{X}_0}, 0) \in MIC(\mathcal{X}_0)$, $(\mathcal{O}_{\mathcal{X}'_0}, 0) \in HIG(\mathcal{X}'_0)$, one obtains the following commutative diagram:

$$\begin{array}{ccc} (\mathcal{O}_{\mathcal{X}_0}, 0) & \xrightarrow{C_{\mathcal{X}_0}} & (\mathcal{O}_{\mathcal{X}'_0}, 0) \\ {\scriptstyle \varsigma(\lambda)^{DR}_*}\big\downarrow \cong & & \cong \big\downarrow{\scriptstyle \varsigma'(\lambda)^{HIG}_*} \\ (\mathcal{O}_{\mathcal{X}_0}, 0) & \xrightarrow{C_{\mathcal{X}_0}} & (\mathcal{O}_{\mathcal{X}'_0}, 0). \end{array}$$

Applying the same theorem further to $f_0 : \mathcal{X}_0 \to \mathcal{M}_0$ one obtains the second commutative diagram:

$$\begin{array}{ccc} (\mathcal{O}_{\mathcal{X}_0}, 0) & \xrightarrow{C_{\mathcal{X}_0}} & (\mathcal{O}_{\mathcal{X}'_0}, 0) \\ {\scriptstyle R^1 f^{DR}_{0*}}\big\downarrow & & \big\downarrow{\scriptstyle R^1 f'^{HIG}_{0*}} \\ (\mathcal{H}^1_{dR}, \nabla) & \xrightarrow{C_{\mathcal{M}_0}} & (E', \theta'). \end{array}$$

It is clear that the isomorphism $\varsigma(\lambda)^{DR}_*$ (resp. $\varsigma'(\lambda)^{HIG}_*$) induces via the direct image functor $R^1 f_{0*}$ the isomorphism $\varsigma(\lambda)^{DR}_* : (\mathcal{H}^1_{dR}, \nabla) \to (\mathcal{H}^1_{dR}, \nabla)$ (resp. the isomorphism $\varsigma'(\lambda)^{HIG}_* : (E', \theta') \to (E', \theta')$). The above two commutative diagrams yield the following commutative diagram:

$$\begin{array}{ccc} (\mathcal{H}^1_{dR}, \nabla) & \xrightarrow{C_{\mathcal{M}_0}} & (E', \theta') \\ {\scriptstyle \varsigma(\lambda)^{DR}_*}\big\downarrow \cong & & \cong \big\downarrow{\scriptstyle \varsigma'(\lambda)^{HIG}_*} \\ (\mathcal{H}^1_{dR}, \nabla) & \xrightarrow{C_{\mathcal{M}_0}} & (E', \theta'). \end{array}$$

Since the action of $\lambda \in \mathcal{O}_{LK}$ on $(\mathcal{H}^1_{dR}, \nabla)$ (resp. (E', θ')) is given by $\varsigma(\lambda)^{DR}_*$ (resp. $\varsigma'(\lambda)^{HIG}_*$), one then has

$$\begin{aligned} \varsigma'(\lambda)\big(C_{\mathcal{M}_0}(\mathcal{H}^1_{dR,\phi}, \nabla_\phi)\big) &= C_{\mathcal{M}_0}\big(\varsigma(\lambda)(\mathcal{H}^1_{dR,\phi}, \nabla_\phi)\big) \\ &= C_{\mathcal{M}_0}\big(\phi(\lambda)(\mathcal{H}^1_{dR,\phi}, \nabla_\phi)\big) \\ &= \phi(\lambda) C_{\mathcal{M}_0}(\mathcal{H}^1_{dR,\phi}, \nabla_\phi), \end{aligned}$$

which implies that $C_{\mathcal{M}_0}(\mathcal{H}^1_{dR,\phi}, \nabla_\phi) = (E'_\phi, \theta'_\phi)$. □

Let $0 = F^2_{\mathrm{con}} \subset F^1_{\mathrm{con}} \subset F^0_{\mathrm{con}} = \mathcal{H}^1_{dR}$ be the conjugate filtration of $\mathcal{H}^1_{dR}$, which is flat with respect to the Gauss–Manin connection (see [13], §3). For a subbundle $W \subset \mathcal{H}^1_{dR}$ we put $Gr_{F_{\mathrm{con}}}(W) = \bigoplus_{q=0}^{1} \frac{W \cap F^q_{\mathrm{con}}}{W \cap F^{q+1}_{\mathrm{con}}}$. The p-curvature ψ_∇ of ∇ defines the F-Higgs bundle

$$\psi_\nabla : Gr_{F_{\mathrm{con}}}(\mathscr{H}^1_{dR}) \to Gr_{F_{\mathrm{con}}}(\mathscr{H}^1_{dR}) \otimes F^*_{\mathscr{M}_0}\Omega_{\mathscr{M}_0}.$$

As a reminder to the reader, we recall the definition of the F-Higgs bundle: an F-Higgs bundle over a base C, which is defined over $\mathbb{F}$, is a pair (E, θ) where E is a vector bundle over C, and θ is a bundle morphism $E \to E \otimes F^*_C\Omega_C$ with the integral property $\theta \wedge \theta = 0$. The following lemma is a simple consequence of Katz's p-curvature formula ([14], Theorem 3.2), and it is also true in a general context.

Lemma 7.2. *Let $W \subset \mathscr{H}^1_{dR}$ be a subbundle preserved by the Gauss–Manin connection ∇. Then the F-Higgs subbundle $\big(Gr_{F_{\mathrm{con}}}(W), \psi_\nabla|_{Gr_{F_{\mathrm{con}}}(W)}\big)$ defines a Higgs subbundle of (E', θ') by the Cartier descent.*

Proof. Since ∇ preserves F_{con}, it induces the connection $Gr_{F_{\mathrm{con}}}\nabla$ on $Gr_{F_{\mathrm{con}}}(\mathscr{H}^1_{dR})$ by taking grading. The operation of the p-curvature on this connection commutes with taking the grading. It follows that $\psi_{Gr_{F_{\mathrm{con}}}\nabla} = Gr_{F_{\mathrm{con}}}\psi_\nabla$, which is the zero map. The relative Cartier isomorphism gives the isomorphism $\big(Gr_{F_{\mathrm{con}}}(\mathscr{H}^1_{dR}), Gr_{F_{\mathrm{con}}}\nabla\big) \cong (\mathscr{F}^*_{\mathscr{M}_0}E', \nabla^{\mathrm{can}})$, where ∇^{can} is the canonical connection by the Cartier descent. Taking this isomorphism for granted, we see that the inclusion $Gr_{F_{\mathrm{con}}}(W) \subset Gr_{F_{\mathrm{con}}}(\mathscr{H}^1_{dR})$ descends to the inclusion $F \subset E'$. In other words, $\mathscr{F}^*_{\mathscr{M}_0}F$ is isomorphic to $Gr_{F_{\mathrm{con}}}(W)$ via the relative inverse Cartier isomorphism. By Katz's formula (see [14]), the F-Higgs bundle $\big(Gr_{F_{\mathrm{con}}}(\mathscr{H}^1_{dR}), \psi_\nabla\big)$ descends to the Higgs bundle (E', θ'). Thus the F-Higgs subbundle $\big(Gr_{F_{\mathrm{con}}}(W), \psi_\nabla|_{Gr_{F_{\mathrm{con}}}(W)}\big)$ descends to the Higgs subbundle $(F, \theta'|_F)$ of (E', θ'). □

For simplicity the above resulting Higgs subbundle is called the *Cartier descent* of $Gr_{F_{\mathrm{con}}}(W, \nabla|_W)$.

Theorem 7.3. *For each $\phi \in \Phi$, the following statements hold*:

(1) $C_{\mathscr{M}_0}(\mathscr{H}^1_{dR,\sigma\phi}, \nabla_{\sigma\phi}) = \pi^*(E_\phi, \theta_\phi)$.

(2) $C_{\mathscr{M}_0}(\mathscr{H}^1_{dR,\phi}, \nabla_\phi)$ *is the Cartier descent of* $Gr_{F_{\mathrm{con}}}(\mathscr{H}^1_{dR,\phi}, \nabla_\phi)$.

The similar statements hold for the bar counterpart.

Proof. By Proposition 7.1 it is equivalent to show the identification $(E_\phi)' = E_{\sigma\phi}$. For this we take $\lambda \in \mathscr{O}_{LK}$. Then

$$\varsigma'(\lambda)\big((E_\phi)'\big) = \big(\varsigma(\lambda)(E_\phi)\big)' = \big(\phi(\lambda)E_\phi\big)' = \phi(\lambda)^p(E_\phi)' = (\sigma\phi)(\lambda)(E_\phi)'.$$

This proves the identification and therefore the first part of the theorem. Because $\theta'|_{E'_\phi} = \theta'_\phi$, the second part is a consequence of Proposition 7.1 and Lemma 7.2. □

By this theorem we have a better understanding of the Higgs subbundles in Proposition 6.6 (ii) under the Frobenius pull-backs. Now assume that $d = [F_{\mathfrak{p}} : \mathbb{Q}_p] > 1$ and write $\mathrm{Hom}_{\mathbb{Q}_p}(L_{\mathfrak{p}}, \overline{\mathbb{Q}}_p) = \{\phi_1, \ldots, \phi_d, \phi^*_1, \ldots, \phi^*_d\}$ as in Proposition 5.1. We deduce the following

Corollary 7.4. *For $2 \leqq i \leqq d$, one has $F_{\mathscr{M}_0}^{*d-i+1} E_{\phi_i} = \mathscr{H}^1_{dR,\phi_1^*}$. Moreover under this identification, the Harder–Narasimhan filtration on $F_{\mathscr{M}_0}^{*d-i+1} E_{\phi_i}$ coincides with the Hodge filtration on $\mathscr{H}^1_{dR,\phi_1^*}$, which is induced from $\mathscr{H}^1_{dR}$.*

Proof. We look at the $i = d$ case first. Since $\big((E_{\phi_d})', 0\big) \in HIG(\mathscr{M}_0')$ is of trivial Higgs field,

$$C^{-1}_{\mathscr{M}_0}\big((E_{\phi_d})', 0\big) = \big(\mathscr{F}^*_{\mathscr{M}_0}(E_{\phi_d})', \nabla^{\mathrm{can}}\big).$$

Because $F_{\mathscr{M}_0} = \pi \circ \mathscr{F}_{\mathscr{M}_0}$, we have $\mathscr{F}^*_{\mathscr{M}_0}(E_{\phi_d})' = F^*_{\mathscr{M}_0} E_{\phi_d}$. On the other hand, the above theorem says that

$$C^{-1}_{\mathscr{M}_0}\big((E_{\phi_d})', 0\big) = C^{-1}_{\mathscr{M}_0}(E'_{\sigma\phi_d}, 0) = C^{-1}_{\mathscr{M}_0}(E'_{\phi_1^*}, 0) = (\mathscr{H}^1_{dR,\phi_1^*}, \nabla_{\phi_1^*}).$$

It follows that $(F^*_{\mathscr{M}_0} E_{\phi_d}, \nabla^{\mathrm{can}}) = (\mathscr{H}^1_{dR,\phi_1^*}, \nabla_{\phi_1^*})$. Forgetting the connections, we obtain the result for $i = d$. For $2 \leqq i \leqq d-1$, the same argument for E_{ϕ_i} implies that $F^*_{\mathscr{M}_0} E_{\phi_i} = \mathscr{H}^1_{dR,\phi_{i+1}}$. As $i+1 \leqq d$, one must have $\mathscr{H}^1_{dR,\phi_{i+1}} = E_{\phi_{i+1}}$. By iterating the above arguments, one obtains the results for all i's from the $i = d$ case. From the second part of the proof of Proposition 6.6, we have known that $F_{\mathscr{M}_0}^{*d-i+1} E_{\phi_i}$ is not semistable. It suffices to show that the Hodge filtration on $\mathscr{H}^1_{dR,\phi_1^*} = F^*_{\mathscr{M}_0} E_{\phi_d}$ is the Harder–Narasimhan filtration. Since the Harder–Narasimhan filtration is unique, it is equivalent to show that the sub line bundle $E^{1,0}_{\phi_1^*} \subset \mathscr{H}^1_{dR,\phi_1^*}$ is of maximal degree. By Theorem 6.2, the maximal degree does not exceed $g-1$, which is exactly the degree of $E^{1,0}_{\phi_1^*}$. This completes the proof. □

Taking the grading of $(F_{\mathscr{M}_0}^{*d-i+1} E_{\phi_i}, \nabla^{\mathrm{can}})$ with respect to the Harder–Narasimhan filtration, one obtains a Higgs bundle different from the original Higgs subbundle $(E_{\phi_i}, 0)$ of unitary type. By the above result, the new Higgs bundle is exactly one of the Higgs subbundles of unformizing type. Such a phenomenon might be quite general for Higgs bundles over compact Shimura varieties in char p.

References

[1] *H. Carayol*, Sur la mauvaise réduction des courbes de Shimura, Comp. Math. **59** (1986), no. 2, 151–230.

[2] *P. Deligne*, Travaux de Shimura, in: Séminaire Bourbaki, Lecture Notes in Math. **244**, Springer, Berlin (1972), 123–165.

[3] *P. Deligne*, Variétés de Shimura: Interpretation modulaire, et techniques de construction de modèles canoniques, Proceedings of Symposia in Pure Mathematics **33** (1979), part 2, 247–290.

[4] *C. Deninger*, Representations attached to vector bundles on curves over finite and p-adic fields, a comparison, preprint 2009, http://arxiv.org/abs/0903.2961

[5] *C. Deninger* and *A. Werner*, Vector bundles on p-adic curves and parallel transport, Ann. Scient. Éc. Norm. Sup. **38** (2005), 553–597.

[6] *T. Ekedahl* and *G. van der Geer*, Cycles representing the top Chern class of the Hodge bundle on the moduli space of abelian varieties, Duke Math. J. **129** (2005), no. 1, 187–199.

[7] *C. Faber* and *E. Looijenga*, Moduli of curves and abelian varieties, The Dutch Intercity Seminar on Moduli, Aspects of Mathematics **E33**, Vieweg, Braunschweig 1999.

[8] *G. Faltings*, A p-adic Simpson correspondence, Advances in Math. **198** (2005), 847–862.

[9] *E. Z. Goren* and *F. Oort*, Stratifications of Hilbert modular varieties, J. Algebraic Geom. **9** (2000), no. 1, 111–154.

[10] *A. Grothendieck*, Groupes de Barsotti-Tate et cristaux de Dieudonné, Séminaire de Mathématiques Supérieures **45**, Presses de l'Université de Montréal, 1974.

[11] *K. Joshi*, *S. Ramanan*, *Z. Xia* and *J. Yu*, On vector bundles destabilized by Frobenius pull-back, Compos. Math. **142** (2006), no. 3, 616–630.
[12] *P. Kassaei*, $\mathscr{P}$-adic modular forms over Shimura curves over totally real fields, Compos. Math. **140** (2004), no. 2, 359–395.
[13] *N. Katz*, Nilpotent connections and the monodromy theorem: Applications of a result of Turrittin, Inst. Hautes Études Sci. Publ. Math. **39** (1970), 175–232.
[14] *N. Katz*, Algebraic solutions of differential equations (p-curvature and the Hodge filtration), Invent. Math. **18** (1972), 1–118.
[15] *N. Katz*, Slope filtration of F-crystals, Astérisque **63** (1979), 113–164.
[16] *N. Katz*, Serre-Tate local moduli, in: Surfacess algébriques, Sém. Géom. Alg. Orsay 1976–78, Lecture Notes in Math. **868**, Springer-Verlag, New York (1981), 138–202.
[17] *H. Lange* and *U. Stuhler*, Vektorbündel auf Kurven und Darstellungen der algebraischen Fundamentalgruppe, Math. Z. **156** (1977), 73–83.
[18] *A. Langer*, Semistable sheaves in positive characteristic, Annals of Math. **159** (2004), no. 1, 251–276.
[19] *S. G. Langton*, Valuative criteria for families of vector bundles on algebraic varieties, Annals of Math. **101** (1975), no. 1, pp. 88–110.
[20] *Q. Liu*, Algebraic geometry and arithmetic curves, Oxf. Grad. Texts Math. **6**, Oxford University Press, 2002.
[21] *J. Lu*, *M. Sheng* and *K. Zuo*, An Arakelov inequality in characteristic p and upper bound of p-rank zero locus, J. Number Theory **129** (2009), 3029–3045.
[22] *D. Mumford*, A note of Shimura's paper "Discontinuous groups and abelian varieties", Math. Ann. **181** (1969), 345–351.
[23] *R. Noot*, On Mumford's families of abelian varieties, J. Pure Appl. Algebra **157** (2001), no. 1, 87–106.
[24] *P. Norman* and *F. Oort*, Moduli of abelian varieties, Annals of Math. **112** (1980), 413–439.
[25] *A. Ogus* and *V. Vologodsky*, Nonabelian Hodge theory in characteristic p, Publ. Math. Inst. Hautes études Sci. **106** (2007), 1–138.
[26] *H. Reimann*, The semi-simple zeta function of quaternionic Shimura varieties, Lecture Notes in Math. **1657**, Springer, Berlin 1997.
[27] *N. Shepherd-Barron*, Semi-stability and reduction mod p, Topology **37** (1998), no. 3, 659–664.
[28] *C. Simpson*, Higgs bundles and local systems, Publ. Math. Inst. Hautes étud. Sci. **75** (1992), 5–95.
[29] *X. T. Sun*, Remarks on semistability of G-bundles in positive characteristic, Compos. Math. **119** (1999), 41–52.
[30] *E. Viehweg* and *K. Zuo*, A characterization of certain Shimura curves in the moduli stack of abelian varieties, J. Differential Geom. **66** (2004), no. 2, 233–287.
[31] *T. Zink*, Cartiertheorie kommutativer formaler Gruppen (German), with English, French and Russian summaries, Teubner Texts in Mathematics **68**, BSB B. G. Teubner Verlagsgesellschaft, Leipzig 1984.

Institut für Mathematik, Johannes Gutenberg-Universität Mainz, 55099 Mainz, Germany
e-mail: sheng@uni-mainz.de

Department of Mathematics, Sichuan University, 610000 Chengdu, P. R. China
e-mail: jjzhang@scu.edu.cn

Institut für Mathematik, Johannes Gutenberg-Universität Mainz, 55099 Mainz, Germany
e-mail: zuok@uni-mainz.de

Eingegangen am 1. August 2010, in revidierter Fassung 25. März 2011