# Algebra & Number Theory

Volume 14
2020
No. 5

msp

# Algebra & Number Theory

msp.org/ant



**EDITORS**

MANAGING EDITOR

Bjorn Poonen
Massachusetts Institute of Technology
Cambridge, USA

EDITORIAL BOARD CHAIR

David Eisenbud
University of California
Berkeley, USA

BOARD OF EDITORS

| | | | |
|---|---|---|---|
| Bhargav Bhatt | University of Michigan, USA | Martin Olsson | University of California, Berkeley, USA |
| Richard E. Borcherds | University of California, Berkeley, USA | Raman Parimala | Emory University, USA |
| Antoine Chambert-Loir | Université Paris-Diderot, France | Jonathan Pila | University of Oxford, UK |
| J-L. Colliot-Thélène | CNRS, Université Paris-Sud, France | Irena Peeva | Cornell University, USA |
| Brian D. Conrad | Stanford University, USA | Anand Pillay | University of Notre Dame, USA |
| Samit Dasgupta | Duke University, USA | Michael Rapoport | Universität Bonn, Germany |
| Hélène Esnault | Freie Universität Berlin, Germany | Victor Reiner | University of Minnesota, USA |
| Gavril Farkas | Humboldt Universität zu Berlin, Germany | Peter Sarnak | Princeton University, USA |
| Hubert Flenner | Ruhr-Universität, Germany | Michael Singer | North Carolina State University, USA |
| Sergey Fomin | University of Michigan, USA | Christopher Skinner | Princeton University, USA |
| Edward Frenkel | University of California, Berkeley, USA | Vasudevan Srinivas | Tata Inst. of Fund. Research, India |
| Wee Teck Gan | National University of Singapore | J. Toby Stafford | University of Michigan, USA |
| Andrew Granville | Université de Montréal, Canada | Shunsuke Takagi | University of Tokyo, Japan |
| Ben J. Green | University of Oxford, UK | Pham Huu Tiep | University of Arizona, USA |
| Joseph Gubeladze | San Francisco State University, USA | Ravi Vakil | Stanford University, USA |
| Christopher Hacon | University of Utah, USA | Michel van den Bergh | Hasselt University, Belgium |
| Roger Heath-Brown | Oxford University, UK | Akshay Venkatesh | Institute for Advanced Study, USA |
| János Kollár | Princeton University, USA | Marie-France Vignéras | Université Paris VII, France |
| Philippe Michel | École Polytechnique Fédérale de Lausanne | Melanie Matchett Wood | University of California, Berkeley, USA |
| Susan Montgomery | University of Southern California, USA | Shou-Wu Zhang | Princeton University, USA |
| Shigefumi Mori | RIMS, Kyoto University, Japan | | |

**PRODUCTION**

production@msp.org
Silvio Levy, Scientific Editor



See inside back cover or msp.org/ant for submission instructions.

The subscription price for 2020 is US $415/year for the electronic version, and $620/year (+$60, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP.

Algebra & Number Theory (ISSN 1944-7833 electronic, 1937-0652 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840 is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

ANT peer review and production are managed by EditFLOW® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/



© 2020 Mathematical Sciences Publishers

msp

ALGEBRA AND NUMBER THEORY 14:5 (2020)
https://doi.org/10.2140/ant.2020.14.1055

# The universal family of semistable $p$-adic Galois representations

Urs Hartl and Eugen Hellmann

Let $K$ be a finite field extension of $\mathbb{Q}_p$ and let $\mathscr{G}_K$ be its absolute Galois group. We construct the universal family of filtered $(\varphi, N)$-modules, or (more generally) the universal family of $(\varphi, N)$-modules with a Hodge–Pink lattice, and study its geometric properties. Building on this, we construct the universal family of semistable $\mathscr{G}_K$-representations in $\mathbb{Q}_p$-algebras. All these universal families are parametrized by moduli spaces which are Artin stacks in schemes or in adic spaces locally of finite type over $\mathbb{Q}_p$ in the sense of Huber. This has conjectural applications to the $p$-adic local Langlands program.



## 1. Introduction

Let $K$ be a finite field extension of $\mathbb{Q}_p$. The emerging $p$-adic local Langlands program wants to relate on the one hand certain continuous representations of the absolute Galois group $\mathscr{G}_K = \mathrm{Gal}(\overline{K}/K)$ of $K$ on $n$-dimensional $L$-vector spaces for another $p$-adic field $L$, and on the other hand topologically irreducible admissible representations of $\mathrm{GL}_n(K)$ on finite-dimensional $L$-Banach spaces in the sense of [Schneider and Teitelbaum 2006]. One fundamental difference to the case where $L$ is an $\ell$-adic field with $\ell \neq p$ is that the $\ell$-adic local Langlands correspondence is a bijection of merely discrete sets. In the $p$-adic case the representations vary in families. So one may even speculate about a continuous or analytic correspondence. At present not even a conjectural formulation of the $p$-adic local Langlands correspondence purely in local terms is known. One of the main tools in the $p$-adic Langlands program is

*MSC2010:* primary 11S20; secondary 11F80, 13A35.
*Keywords:* p-adic Galois representations, crystalline representations, semistable representations, moduli spaces, filtered modules.

to consider families of representations that admit a dense set of points, where the representations "come from a global set-up", as in [Caraiani et al. 2016] for example. Hence a good understanding of these arithmetic families of $p$-adic Galois representations of $\mathcal{G}_K$ seems to be crucial. This understanding is our aim in the present article: we develop notions of $p$-adic families of $p$-adic Hodge structures (such as filtered $(\varphi, N)$-modules) and $p$-adic Galois representations and study the relation between these two.

The study of such families was begun in [Kisin 2006; 2008; Pappas and Rapoport 2009] and in [Hellmann 2013], where a universal family of filtered $\varphi$-modules was constructed and, building on this, a universal family of crystalline representations with Hodge–Tate weights in $\{0, 1\}$. The approach is based on Kisin's integral $p$-adic Hodge theory cf. [Kisin 2006].

In the present article we generalize these results in two directions. First we consider more general families of $p$-adic Hodge-structure, namely families of $(\varphi, N)$-modules together with a so called *Hodge–Pink lattice*. The inspiration to work with Hodge–Pink lattices instead of filtrations is taken from the analogous theory over function fields; see [Pink 1997; Genestier and Lafforgue 2011; Hartl 2011]. It was already applied to Kisin's integral $p$-adic Hodge theory by Genestier and Lafforgue [2012] in the absolute case for $\varphi$-modules over $\mathbb{Q}_p$.

Second we generalize [Hellmann 2013] to the case of semistable representations. In doing so we correct some mistakes made in [loc. cit.]. The generalization to families with more general Hodge–Tate weights than those in [loc. cit.] (where the weights are assumed to be in $\{0, 1\}$) gives another good reason to work with families of Hodge–Pink lattices: Kisin's theory does not describe $\mathcal{G}_K$-stable $\mathbb{Z}_p$-lattices in a crystalline (or semistable) $\mathcal{G}_K$-representation but all $\mathcal{G}_{K_\infty}$-stable $\mathbb{Z}_p$-lattices, where $K_\infty$ is a certain Kummer extension of $K$ appearing in [Kisin 2006]. The $(\varphi, N)$-modules with a Hodge–Pink lattice correspond to certain $\mathcal{G}_{K_\infty}$ representations and we describe their moduli space (or stack). This stack turns out to be a vector bundle over a space of filtered $(\varphi, N)$-modules. The original space of filtered $(\varphi, N)$-modules (corresponding to $\mathcal{G}_K$ rather than $\mathcal{G}_{K_\infty}$-representations) can be recovered as a section defined by a certain transversality condition in this vector bundle. Moreover, we consider (following [Pappas and Rapoport 2009]) a stack of integral $p$-adic Hodge-structures and a period morphism to the moduli stack of $(\varphi, N)$-modules with a Hodge–Pink lattice and describe its image. Once again, this only works using the more general framework of Hodge–Pink lattices.

The introduction of Hodge–Pink lattices rather than filtrations shows new and interesting phenomena: similarly to the case of filtrations one can define weights of a Hodge–Pink lattice. However, these weights can jump within a family! Whereas for families of $\mathcal{G}_K$-representations the Hodge–Tate weights should vary continuously. On the Galois side there is an explanation of this behavior as follows: there is no notion of Hodge–Tate weights for representations of $\mathcal{G}_{K_\infty}$, but only for representations of $\mathcal{G}_K$.

We can define a notion of weak admissibility for $(\varphi, N)$-modules with Hodge–Pink lattice and show that being weakly admissible is an open condition in the set-up of adic spaces generalizing the corresponding result for filtered $\varphi$-modules in [Hellmann 2013]. Following the method of [Kisin 2006; Hellmann 2013] we further cut out an open subspace over which an integral structure for the $(\varphi, N)$-modules with Hodge–Pink lattice exists and an open subspace over which a family of $\mathcal{G}_{K_\infty} = \mathrm{Gal}(\overline{K}/K_\infty)$-representation

exists. If we restrict ourselves to the subspace of filtered $(\varphi, N)$-modules one can promote this family of $\mathscr{G}_{K_\infty}$-representations to the universal family of semistable $\mathscr{G}_K$-representations.

We describe our results in more detail. Let $K$ be a finite extension of $\mathbb{Q}_p$ with absolute Galois group $\mathscr{G}_K$ and maximal unramified subextension $K_0$. Let $\mathrm{Frob}_p$ be the $p$-Frobenius on $K_0$. We consider *families of* $(\varphi, N)$*-modules* over $\mathbb{Q}_p$-schemes $X$, that is finite locally free $\mathcal{O}_X \otimes_{\mathbb{Q}_p} K_0$-modules $D$ together with a $\varphi := \mathrm{id} \otimes \mathrm{Frob}_p$-linear automorphism $\Phi$ and a linear monodromy operator $N : D \to D$ satisfying the usual relation $N\Phi = p\Phi N$. Choosing locally on $X$ a basis of $D$ and considering $\Phi \in \mathrm{GL}_d(\mathcal{O}_X \otimes_{\mathbb{Q}_p} K_0)$ and $N \in \mathrm{Mat}_{d\times d}(\mathcal{O}_X \otimes_{\mathbb{Q}_p} K_0)$ as matrices, the condition $N\Phi = p\Phi N$ cuts out a closed subscheme $P_{K_0,d} \subset \mathrm{Res}_{K_0/\mathbb{Q}_p} \mathrm{GL}_d \times_{\mathbb{Q}_p} \mathrm{Res}_{K_0/\mathbb{Q}_p} \mathrm{Mat}_{d\times d}$. We can describe the geometry of this scheme as follows.

**Theorem 3.2.** *The scheme $P_{K_0,d}$ is equidimensional of dimension $[K_0 : \mathbb{Q}_p]\, d^2$. It is reduced, Cohen–Macaulay and generically smooth over $\mathbb{Q}_p$. Its irreducible components are indexed by the possible Jordan types of the* (*necessarily nilpotent*) *monodromy operator $N$.*

Further we consider families of $(\varphi, N)$-modules $D$ with a filtration $\mathcal{F}^\bullet$ on $D \otimes_{K_0} K$ and more generally families of $(\varphi, N)$-modules with a Hodge–Pink lattice $\mathfrak{q}$; see Definition 2.5 for the precise definitions. Given a cocharacter $\mu$ of the algebraic group $\mathrm{Res}_{K/\mathbb{Q}_p} \mathrm{GL}_{d,K}$ (or more precisely a cocharacter of the Weil restriction of the diagonal torus which is dominant with respect to the Weil restriction of the upper triangular matrices) we define the notions of a filtration $\mathcal{F}^\bullet$ and a Hodge–Pink lattice with *constant Hodge polygon equal to* $\mu$, and the notion of *boundedness by* $\mu$ for a Hodge–Pink lattice $\mathfrak{q}$. Associated with $\mu$ is a reflex field $E_\mu$ which is a finite extension of $\mathbb{Q}_p$.

**Theorem 3.6.** (a) *The stack $\mathscr{H}_{\varphi,N,\preceq\mu}$ parametrizing rank $d$ families of $(\varphi, N)$-modules with Hodge–Pink lattice bounded by $\mu$ on the category of $E_\mu$-schemes is an Artin stack. It is equidimensional and generically smooth. Its dimension can be explicitly described in terms of the cocharacter $\mu$ and its irreducible components are indexed by the possible Jordan types of the* (*nilpotent*) *monodromy operator.*

(b) *The stack $\mathscr{H}_{\varphi,N,\mu}$ parametrizing rank $d$ families of $(\varphi, N)$-modules with Hodge–Pink lattice with constant Hodge polygon equal to $\mu$, is an open and dense substack of $\mathscr{H}_{\varphi,N,\preceq\mu}$. Further it is reduced and Cohen–Macaulay. It admits a canonical map to the stack $\mathscr{D}_{\varphi,N,\mu}$ of filtered $(\varphi, N)$-modules with filtration of type $\mu$. This map is representable by a vector bundle.*

If we restrict ourselves to the case of vanishing monodromy, i.e., the case $N = 0$, we cut out a single irreducible component $\mathscr{H}_{\varphi,\preceq\mu} \subset \mathscr{H}_{\varphi,N,\preceq\mu}$ and similarly for the other stacks in the theorem. Following [Hellmann 2013] we consider the above stacks also as stacks on the category of adic spaces locally of finite type over $\mathbb{Q}_p$, i.e., we consider the adification $\mathscr{H}^{\mathrm{ad}}_{\varphi,N,\preceq\mu}$, etc. Passing from $\mathbb{Q}_p$-schemes to adic spaces allows us to generalize Kisin's comparison between filtered $(\varphi, N)$-modules and vector bundles on the open unit disc (together with certain additional structures). To do so we need to fix a uniformizer $\pi$ of $K$ as well as its minimal polynomial $E(u)$ over $K_0$.

**Theorem 4.6.** *For every adic space $X$ locally of finite type over $\mathbb{Q}_p$ there is a natural equivalence of categories between the category of $(\varphi, N)$-modules with Hodge–Pink lattice over $X$ and the category of $(\varphi, N_\nabla)$-modules over $X$, i.e., the category of vector bundles $\mathcal{M}$ on the product of $X$ with the open unit disc $\mathbb{U}$ over $K_0$ together with a semilinear map $\Phi_\mathcal{M} : \mathcal{M} \to \mathcal{M}$ that is an isomorphism away from $X \times \{E(u) = 0\} \subset X \times \mathbb{U}$ and a differential operator $N_\nabla^\mathcal{M}$ satisfying*

$$N_\nabla^\mathcal{M} \circ \Phi_\mathcal{M} \circ \varphi = p \frac{E(u)}{E(0)} \cdot \Phi_\mathcal{M} \circ \varphi \circ N_\nabla^\mathcal{M}.$$

**Theorem 4.9.** *The differential operator $N_\nabla^\mathcal{M}$ defines a canonical meromorphic connection on the vector bundle $\mathcal{M}$. The closed substack $\mathscr{H}^\nabla_{\varphi,N,\mu} \subset \mathscr{H}^{\mathrm{ad}}_{\varphi,N,\mu}$ where this connection is holomorphic coincides with the zero section of the vector bundle $\mathscr{H}^{\mathrm{ad}}_{\varphi,N,\mu} \to \mathscr{D}^{\mathrm{ad}}_{\varphi,N,\mu}$.*

It should be mentioned that the results of Kisin [2006] have a parallel story, earlier developed by Berger [2002], using the cyclotomic extension $K(\varepsilon_n, n \geq 1)$ for a compatible system $(\varepsilon_n)$ of $p^n$-th root of unity, instead of the Kummer extension $K_\infty$. The above results are very much inspired by [loc. cit.].

Similarly to the case of filtered $\varphi$-modules in [Hellmann 2013] there is a notion of weak admissibility for families of $(\varphi, N)$-modules with Hodge–Pink lattice over an adic space. We show that weak admissibility is an open condition.

**Theorem 5.6.** *Let $\mu$ be a cocharacter as above. Then the groupoid*

$$X \mapsto \left\{(D, \Phi, N, \mathfrak{q}) \in \mathscr{H}_{\varphi,N,\preceq\mu}(X) \mid D \otimes \kappa(x) \text{ is weakly admissible for all } x \in X\right\}$$

*is an open substack $\mathscr{H}^{\mathrm{ad,wa}}_{\varphi,N,\preceq\mu}$ of $\mathscr{H}^{\mathrm{ad}}_{\varphi,N,\preceq\mu}$.*

Following the construction in [Hellmann 2013] we construct an open substack $\mathscr{H}^{\mathrm{ad,int}}_{\varphi,N,\preceq\mu} \subset \mathscr{H}^{\mathrm{ad,wa}}_{\varphi,N,\preceq\mu}$ where an integral model for the $(\varphi, N_\nabla)$-module over the open unit disc exists. Here integral means with respect to the ring of integers $W$ in $K_0$. Dealing with Hodge–Pink lattices instead of filtrations makes it possible to generalize the period morphism of [Pappas and Rapoport 2009, §5] beyond the miniscule case. That is, we consider a stack $\widehat{\mathcal{C}}_{\preceq\mu,N,K}$ in the category of formal schemes over $\operatorname{Spf} \mathcal{O}_{E_\mu}$ whose $R$-valued points parameterize tuples $(\mathfrak{M}, \Phi, N)$, where $\mathfrak{M}$ is a finite locally free $(R \otimes_{\mathbb{Z}_p} W)[\![u]\!]$ module, $\Phi$ is a semilinear morphism $\Phi : \mathfrak{M} \to \mathfrak{M}$ which is an isomorphism away from $E(u) = 0$, whose behavior at $E(u)$ is controlled in terms of $\mu$, and $N$ is an endomorphism of $\mathfrak{M}/u\mathfrak{M}$ satisfying $N\Phi = p\Phi N$; see after Remark 6.8 for the precise definition.

Given a $p$-adic formal scheme $\mathcal{X}$ over $\operatorname{Spf} \mathcal{O}_{E_\mu}$ we construct a period morphism

$$\Pi(\mathcal{X}) : \widehat{\mathcal{C}}_{\preceq\mu,N,K}(\mathcal{X}) \to \mathscr{H}^{\mathrm{ad}}_{\varphi,N,\preceq\mu}(\mathcal{X}^{\mathrm{ad}})$$

and the substack $\mathscr{H}^{\mathrm{ad,int}}_{\varphi,N,\preceq\mu}$ will serve as the image of this morphism in the following sense:

**Corollary 6.10.** *Let $X$ be an adic space locally of finite type over the reflex field $E_\mu$ of $\mu$ and let $f : X \to \mathscr{H}^{\mathrm{ad}}_{\varphi,N,\preceq\mu}$ be a morphism defined by $(D, \Phi, N, \mathfrak{q})$. Then $f$ factors over $\mathscr{H}^{\mathrm{ad,int}}_{\varphi,N,\preceq\mu}$ if and only if there exists an fpqc-covering $(U_i \to X)_{i\in I}$ and formal models $\mathcal{U}_i$ of $U_i$ together with $(\mathfrak{M}_i, \Phi_i) \in \widehat{\mathcal{C}}_{\preceq\mu,N,K}(\mathcal{U}_i)$ such that $\Pi(\mathcal{U}_i)(\mathfrak{M}_i, \Phi_i) = (D, \Phi, N, \mathfrak{q})|_{U_i}$.*

Finally we go back to Galois representations. We prove that there is a canonical open subspace $\mathscr{H}^{\mathrm{red,ad,adm}}_{\varphi,N,\preceq\mu}$ of the reduced space underlying $\mathscr{H}^{\mathrm{ad,int}}_{\varphi,N,\preceq\mu}$ which carries a family of $\mathscr{G}_{K_\infty}$-representations. This family is universal in a sense made precise in the body of the article. Roughly this means that a morphism $f : X \to \mathscr{H}^{\mathrm{ad,int}}_{\varphi,N,\preceq\mu}$ defined by some $(\mathfrak{M}, \Phi, N)$ over a formal model $\mathcal{X}$ of $X$ factors over $\mathscr{H}^{\mathrm{ad,adm}}_{\varphi,N,\preceq\mu}$ if and only there exists a family of $\mathscr{G}_{K_\infty}$-representations $\mathcal{E}$ on $X$ such that the $\varphi$-module of $\mathcal{E}$, in the sense of Fontaine, is (up to inverting $p$) given by the $p$-adic completion of $(\mathfrak{M}, \Phi)[1/u]$. For a finite extension $L$ of $E_\mu$, Kisin's theory implies that we have an equality

$$\mathscr{H}^{\mathrm{red,ad,adm}}_{\varphi,N,\preceq\mu}(L) = \mathscr{H}^{\mathrm{ad,int}}_{\varphi,N,\preceq\mu}(L) = \mathscr{H}^{\mathrm{ad,wa}}_{\varphi,N,\preceq\mu}(L)$$

of $L$-valued points.

If we want to promote our family of $\mathscr{G}_{K_\infty}$-representations to a family of $\mathscr{G}_K$-representations, we have to restrict ourselves to filtrations rather than Hodge–Pink lattices. The reason for that is that the meromorphic connection $\nabla$ in Theorem 4.9 above must be holomorphic in this case. In the framework of Berger's work [2002] with the cyclotomic tower, this is in some sense even more apparent: the connection $\nabla$ comes from the derivation of the $\Gamma$-action.

**Theorem 8.15.** *There is an open substack $\mathscr{D}^{\mathrm{ad,adm}}_{\varphi,N,\mu} \subset \mathscr{D}^{\mathrm{ad}}_{\varphi,N,\mu}$ over which there exists a family $\mathcal{E}$ of semistable $\mathscr{G}_K$-representations such that $D_{\mathrm{st}}(\mathcal{E}) = (D, \Phi, N, \mathcal{F}^\bullet)$ is the restriction of the universal family of filtered $(\varphi, N)$-modules on $\mathscr{D}^{\mathrm{ad}}_{\varphi,N,\mu}$ to $\mathscr{D}^{\mathrm{ad,adm}}_{\varphi,N,\mu}$.*

*This family is universal in the following sense: Let $X$ be an adic space locally of finite type over the reflex field $E_\mu$ of $\mu$, and let $\mathcal{E}'$ be a family of semistable $\mathscr{G}_K$-representations on $X$ with constant Hodge polygon equal to $\mu$. Then there is a unique morphism $f : X \to \mathscr{D}^{\mathrm{ad,adm}}_{\varphi,N,\mu}$ such that $\mathcal{E}' \cong f^*\mathcal{E}$ as families of $\mathscr{G}_K$-representations.*

The corresponding result for crystalline $\mathscr{G}_K$-representations with constant Hodge polygon equal to $\mu$, whose moduli space is $\mathscr{D}^{\mathrm{ad,adm}}_{\varphi,\mu}$, is formulated and proved in Corollary 8.16. We finally briefly discuss how these results relate to Kisin's construction of potentially semistable deformation rings [Kisin 2008]. There is a precise relation between our universal family and Kisin's construction discussed in Proposition 8.17. It should be mentioned however, that the spirit of our approach differs from Kisin's: we study families of $p$-adic Hodge-structures (i.e., semilinear algebra data) and then cut out a subspace defining a Galois representation. Kisin starts with families of Galois representations (provided by deformation rings) and then cuts out a crystalline locus. Moreover, his definition of a crystalline family differs from ours: Kisin defines a family to be semistable if its base change to all finite-dimensional $\mathbb{Q}_p$-algebras is semistable. In contrast we aim at giving a definition that is more in the spirit of Fontaine's definition using period rings. In fact, as we needed to correct some mistakes from the last section of [Hellmann 2013], we also changed the definition of crystalline representations from [loc. cit.]: it seems to be a bit messy to deal with the filtration on a sheafified version of $B_{\mathrm{cris}}$, hence we rather use the $\varphi$-modules on the open unit disc as our $p$-adic Hodge structures and define the notion of a semistable representation using the comparison of a vector bundle on the open unit disc and a Galois representation after tensoring with (a relative version of) $B^+_{\mathrm{cris}}$; see Definition 8.2.

***Notations.*** Let $K$ be a finite field extension of the $p$-adic numbers $\mathbb{Q}_p$ and fix an algebraic closure $\overline{K}$ of $K$. We write $\mathbb{C}_p$ for the $p$-adic completion of $\overline{K}$ and let $\mathcal{G}_K = \mathrm{Gal}(\overline{K}/K)$ be the absolute Galois group of $K$. Let $\widetilde{K}$ be the Galois closure of $K$ inside $\overline{K}$. Let $K_0$ be the maximal unramified subfield of $K$ and $W$ its ring of integers. Set $f := [K_0 : \mathbb{Q}_p]$, and let $\mathrm{Frob}_p$ be the Frobenius automorphism of $K_0$ which induces the $p$-power map on the residue field of $K_0$. We fix once and for all a uniformizer $\pi$ of $K$ and its minimal polynomial $E(u) = \mathrm{mipo}_{\pi/K_0}(u) \in W[u]$ over $K_0$. It is an Eisenstein polynomial, and $K = K_0[u]/(E(u))$. We choose a compatible system $\pi_n$ of $p^n$-th roots of $\pi$ in $\overline{K}$ and write $K_\infty$ for the field obtained from $K$ by adjoining all $\pi_n$.

## 2. Families of $(\varphi, N)$-modules with Hodge–Pink lattice

Let $R$ be a $\mathbb{Q}_p$-algebra and consider the endomorphism $\varphi := \mathrm{id}_R \otimes \mathrm{Frob}_p$ of $R \otimes_{\mathbb{Q}_p} K_0$. For an $R \otimes_{\mathbb{Q}_p} K_0$-module $M$ we set $\varphi^* M := M \otimes_{R \otimes_{\mathbb{Q}_p} K_0, \varphi} R \otimes_{\mathbb{Q}_p} K_0$. Similar notation is applied to morphisms between $R \otimes_{\mathbb{Q}_p} K_0$-modules. We let $\varphi^* : M \to \varphi^* M$ be the $\varphi$-semilinear map with $\varphi^*(m) = m \otimes 1$.

We introduce the rings

$$\mathbb{B}_R^+ := \varprojlim_i (R \otimes_{\mathbb{Q}_p} K_0[u])/(E(u)^i) \quad \text{and} \quad \mathbb{B}_R := \mathbb{B}_R^+\Big[\frac{1}{E(u)}\Big].$$

In a certain sense $\mathbb{B}^+_{\mathbb{Q}_p}$ and $\mathbb{B}_{\mathbb{Q}_p}$ are the analogues of Fontaine's rings $\mathbb{B}^+_{\mathrm{dR}}$ and $\mathbb{B}_{\mathrm{dR}}$ in Kisin's theory [2006] of $p$-adic Galois representation. By Cohen's structure theorem [Serre 1979, Theorem II.4.2] the ring $\mathbb{B}^+_{\mathbb{Q}_p} = \varprojlim K_0[u]/(E(u)^i)$ is isomorphic to $K[\![t]\!]$ under a map sending $t$ to $E(u)/E(0)$ (and by Hensel's lemma the lift of the residue field $K$ to a subring of $\mathbb{B}^+_{\mathbb{Q}_p}$ is unique). The rings $\mathbb{B}_R^+$ and $\mathbb{B}_R$ are relative versions over $R$, and are isomorphic to $(R \otimes_{\mathbb{Q}_p} K)[\![t]\!]$ and $(R \otimes_{\mathbb{Q}_p} K)[\![t]\!][1/t]$, respectively. We extend $\varphi$ to $R \otimes_{\mathbb{Q}_p} K_0[u]$ by requiring that $\varphi(u) = u^p$ and we define $\varphi^n(\mathbb{B}_R^+) := \varprojlim_i (R \otimes_{\mathbb{Q}_p} K_0[u])/(\varphi^n(E(u))^i)$. We may also identify $\varphi^n(\mathbb{B}_R^+)$ with

$$\varprojlim (R \otimes_{\mathbb{Q}_p} K(\pi_n)[u])/(1 - (u/\pi_n))^i = (R \otimes_{\mathbb{Q}_p} K(\pi_n))[\![1 - (u/\pi_n)]\!]$$

under the assignment $E(u)/E(0) \mapsto 1 - (u/\pi_n)$; compare [Kisin 2006, (1.1.1)]. We extend these rings to sheaves of rings $\varphi^n(\mathbb{B}_X^+) := \varphi^n(\mathbb{B}^+_{\mathcal{O}_X})$ on $\mathbb{Q}_p$-schemes $X$ or adic spaces $X \in \mathrm{Ad}^{\mathrm{lft}}_{\mathbb{Q}_p}$. Here $\mathrm{Ad}^{\mathrm{lft}}_{\mathbb{Q}_p}$ denotes the category of adic spaces locally of finite type, see [Huber 1994] for example.

**Remark 2.1.** Note that $\varphi^n(\mathbb{B}_R^+)$ is not a subring of $\mathbb{B}_R^+$. If $X = \mathrm{Spa}(R, R^\circ)$ is an affinoid adic space of finite type over $\mathbb{Q}_p$ one should think of $\varphi^n(\mathbb{B}_R^+)$ as the completion of the structure sheaf on $X \times \mathbb{U}$ along the section defined by $\varphi^n(E(u)) \in \mathbb{U}$. Here $\mathbb{U}$ denotes the open unit disc over $K_0$.

**Definition 2.2.** (a) A *$\varphi$-module* $(D, \Phi)$ *over* $R$ consists of a locally free $R \otimes_{\mathbb{Q}_p} K_0$-module $D$ of finite rank, and an $R \otimes_{\mathbb{Q}_p} K_0$-linear isomorphism $\Phi : \varphi^* D \xrightarrow{\sim} D$. A *morphism* $\alpha : (D, \Phi) \to (\widetilde{D}, \widetilde{\Phi})$ *of $\varphi$-modules* is an $R \otimes_{\mathbb{Q}_p} K_0$-homomorphism $\alpha : D \to \widetilde{D}$ with $\alpha \circ \Phi = \widetilde{\Phi} \circ \varphi^* \alpha$.

(b) A *$(\varphi, N)$-module* $(D, \Phi, N)$ *over* $R$ consists of a $\varphi$-module $(D, \Phi)$ over $R$ and an $R \otimes_{\mathbb{Q}_p} K_0$-linear endomorphism $N : D \to D$ satisfying $N \circ \Phi = p \cdot \Phi \circ \varphi^* N$. A *morphism* $\alpha : (D, \Phi, N) \to (\widetilde{D}, \widetilde{\Phi}, \widetilde{N})$ *of $(\varphi, N)$-modules* is a morphism of $\varphi$-modules with $\alpha \circ N = \widetilde{N} \circ \alpha$.

The rank of $D$ over $R \otimes_{\mathbb{Q}_p} K_0$ is called the *rank* of $(D, \Phi)$ or $(D, \Phi, N)$.

Every $\varphi$-module over $R$ can be viewed as a $(\varphi, N)$-module with $N = 0$.

**Lemma 2.3.** (a) *Every $\varphi$-module $(D, \Phi)$ over $R$ is Zariski locally on* Spec $R$ *free over $R \otimes_{\mathbb{Q}_p} K_0$.*

(b) *The endomorphism $N$ of a $(\varphi, N)$-module over $R$ is automatically nilpotent.*

*Proof.* (a) Let $\mathfrak{m} \subset R$ be a maximal ideal. Then $R/\mathfrak{m} \otimes_{\mathbb{Q}_p} K_0$ is a direct product of fields which are transitively permuted by $\mathrm{Gal}(K_0/\mathbb{Q}_p)$. The existence of the isomorphism $\Phi$ implies that $D \otimes_R R/\mathfrak{m}$ is free over $R/\mathfrak{m} \otimes_{\mathbb{Q}_p} K_0$. Now the assertion follows by Nakayama's lemma.

(b) By (a) we may locally on $R$ write $N$ as a matrix with entries in $R \otimes_{\mathbb{Q}_p} K_0$. Set $d := \mathrm{rk}\, D$. If the entries of the $d$-th power $N^d$ lie in $\mathrm{Rad}(0) \otimes_{\mathbb{Q}_p} K_0$, where $\mathrm{Rad}(0) = \bigcap_{\mathfrak{p} \subset R \text{ prime}} \mathfrak{p}$ is the nil-radical, then $N$ is nilpotent. Thus we may check the assertion in $L = \mathrm{Frac}(R/\mathfrak{p})^{\mathrm{alg}}$ for all primes $\mathfrak{p} \subset R$. We replace $R$ by $L$. Then $D = \prod V_\psi$ splits up into a direct product of $d$-dimensional $L$-vector spaces indexed by the embeddings $\psi : K_0 \hookrightarrow L$. For every fixed embedding $\psi$ the $f$-th power $\Phi^f$ restricts to an endomorphism $\Phi_\psi$ of $V_\psi$ satisfying $N\Phi_\psi = p^f \Phi_\psi N$. If $V(\lambda, \Phi_\psi)$ denotes the generalized eigenspace for some $\lambda \in L^\times$, then $N$ maps $V(\lambda, \Phi_\psi)$ to $V(p^f\lambda, \Phi_\psi)$ and hence $N$ is nilpotent, as there are only finitely many nonzero eigenspaces. This implies that $N^d = 0$. □

**Remark 2.4.** If $R$ is even a $K_0$-algebra, we can decompose $R \otimes_{\mathbb{Q}_p} K_0 \cong \prod_{i \in \mathbb{Z}/f\mathbb{Z}} R$ where the $i$-th factor is given by the map $R \otimes_{\mathbb{Q}_p} K_0 \to R$, $a \otimes b \mapsto a\,\mathrm{Frob}_p^{-i}(b)$ for $a \in R$, $b \in K_0$. For a $(\varphi, N)$-module over $R$ we obtain corresponding decompositions $D = \prod_i D_i$ and $\varphi^* D = \prod_i (\varphi^* D)_i$ with $(\varphi^* D)_i = D_{i-1}$, and therefore also $\Phi = (\Phi_i : D_{i-1} \xrightarrow{\sim} D_i)_i$ and $N = (N_i : D_i \to D_i)_i$ with $p\,\Phi_i \circ N_{i-1} = N_i \circ \Phi_i$, because $(\varphi^* N)_i = N_{i-1}$. If we set

$$\Psi_i := \Phi_i \circ \cdots \circ \Phi_1 = (\Phi \circ \varphi^*\Phi \circ \cdots \circ \varphi^{(i-1)*}\Phi)_i : D_0 = (\varphi^{i*} D)_i \xrightarrow{\sim} D_i,$$

then $p^i\, \Psi_i \circ N_0 = N_i \circ \Psi_i$ for all $i$, and $\Psi_f = (\Phi^f)_0$. There is an isomorphism of $(\varphi, N)$-modules over $R$

$$(\mathrm{id}_{D_0}, \Psi_1, \ldots, \Psi_{f-1}) : \Big(\prod_i D_0,\ \big((\Phi^f)_0,\ \mathrm{id}_{D_0}, \ldots,\ \mathrm{id}_{D_0}\big),\ (p^i N_0)_i\Big) \xrightarrow{\sim} \Big(\prod_i D_i,\ (\Phi_i)_i,\ (N_i)_i\Big). \quad (2\text{-}1)$$

Thus $(D, \Phi, N)$ is uniquely determined by $(D_0, (\Phi^f)_0, N_0)$ satisfying $p^f (\Phi^f)_0 \circ N_0 = N_0 \circ (\Phi^f)_0$. Further note that under this isomorphism $\Phi^f$ on $(D, \Phi, N)$ corresponds to $\big((\Phi^f)_0, \ldots, (\Phi^f)_0\big)$ on the left-hand side.

**Definition 2.5.** (a) A *$K$-filtered $(\varphi, N)$-module* $(D, \Phi, N, \mathcal{F}^\bullet)$ *over* $R$ consists of a $(\varphi, N)$-module $(D, \Phi, N)$ over $R$ together with a decreasing separated and exhaustive $\mathbb{Z}$-filtration $\mathcal{F}^\bullet$ on $D_K := D \otimes_{K_0} K$ by $R \otimes_{\mathbb{Q}_p} K$-submodules such that $\mathrm{gr}^i_{\mathcal{F}} D_K := \mathcal{F}^i D_K / \mathcal{F}^{i+1} D_K$ is locally free as an $R$-module for all $i$. A *morphism* $\alpha : (D, \Phi, N, \mathcal{F}^\bullet) \to (\widetilde{D}, \widetilde{\Phi}, \widetilde{N}, \widetilde{\mathcal{F}}^\bullet)$ is a morphism of $(\varphi, N)$-modules with $(\alpha \otimes \mathrm{id})(\mathcal{F}^i D_K) \subset \widetilde{\mathcal{F}}^i \widetilde{D}_K$.

(b) A *$(\varphi, N)$-module with Hodge–Pink lattice* $(D, \Phi, N, \mathfrak{q})$ *over* $R$ consists of a $(\varphi, N)$-module $(D, \Phi, N)$ over $R$ together with a *$\mathbb{B}_R^+$-lattice* $\mathfrak{q} \subset D \otimes_{R \otimes K_0} \mathbb{B}_R$. This means that $\mathfrak{q}$ is a finitely generated $\mathbb{B}_R^+$-submodule,

which is a direct summand as $R$-module satisfying $\mathbb{B}_R \cdot \mathfrak{q} = D \otimes_{R\otimes K_0} \mathbb{B}_R$. We call $\mathfrak{q}$ the *Hodge–Pink lattice* of $(D, \Phi, N, \mathfrak{q})$. A *morphism* $\alpha : (D, \Phi, N, \mathfrak{q}) \to (\widetilde{D}, \widetilde{\Phi}, \widetilde{N}, \tilde{\mathfrak{q}})$ is a morphism of $(\varphi, N)$-modules with $(\alpha \otimes \mathrm{id})(\mathfrak{q}) \subset \tilde{\mathfrak{q}}$.

**Remark 2.6.** Note that the graded pieces $\mathrm{gr}^i_{\mathcal{F}} D_K$ in (a) are $R \otimes_{\mathbb{Q}_p} K$-modules that are locally on $\mathrm{Spec}(R \otimes_{\mathbb{Q}_p} K)$ free, but not necessarily of the same rank. Hence they are not locally on $\mathrm{Spec}\, R$ free as $R \otimes_{\mathbb{Q}_p} K$-modules. However, they are locally on $\mathrm{Spec}\, R$ free as $R$-modules.

For every $(\varphi, N)$-module with Hodge–Pink lattice $(D, \Phi, N, \mathfrak{q})$ over $R$ we also consider the tautological $\mathbb{B}^+_R$-lattice $\mathfrak{p} := D \otimes_{R\otimes K_0} \mathbb{B}^+_R$.

**Lemma 2.7.** *Let $\mathfrak{q} \subset D \otimes_{R\otimes K_0} \mathbb{B}_R$ be a $\mathbb{B}^+_R$-submodule. Then $\mathfrak{q}$ is a $\mathbb{B}^+_R$-lattice if and only if $E(u)^n\mathfrak{p} \subset \mathfrak{q} \subset E(u)^{-m}\mathfrak{p}$ for all $n, m \gg 0$ and for any (some) such $n, m$ the quotients $E(u)^{-m}\mathfrak{p}/\mathfrak{q}$ and $\mathfrak{q}/E(u)^n\mathfrak{p}$ are finite locally free $R$-modules.*

*If this is the case then étale locally on* $\mathrm{Spec}\, R$ *the $\mathbb{B}^+_R$-module $\mathfrak{q}$ is free of the same rank as $\mathfrak{p}$.*

*Proof.* The assertion $E(u)^n\mathfrak{p} \subset \mathfrak{q} \subset E(u)^{-m}\mathfrak{p}$ for all $n, m \gg 0$ is equivalent to $\mathbb{B}_R \cdot \mathfrak{q} = D \otimes_{R\otimes K_0} \mathbb{B}_R$ when $\mathfrak{q}$ is finitely generated. Consider such $n, m$. If $\mathfrak{q}$ is a $\mathbb{B}^+_R$-lattice, hence a direct summand of $D \otimes_{R\otimes K_0} \mathbb{B}_R$ there is an $R$-linear section $s$ of the projection $\mathrm{pr} : D \otimes_{R\otimes K_0} \mathbb{B}_R \twoheadrightarrow (D \otimes_{R\otimes K_0} \mathbb{B}_R)/\mathfrak{q}$. The composition of this section with the inclusion $E(u)^{-m}\mathfrak{p}/\mathfrak{q} \hookrightarrow (D \otimes_{R\otimes K_0} \mathbb{B}_R)/\mathfrak{q}$ factors through $E(u)^{-m}\mathfrak{p}$: Indeed, for $x \in E(u)^{-m}\mathfrak{p}/\mathfrak{q}$ the condition $\mathrm{pr}(s(x)) = x$ means that there exists $x' \in \mathfrak{q}$ such that $s(x) = x + x' \in E(u)^{-m}\mathfrak{p} + \mathfrak{q} = E(u)^{-m}\mathfrak{p}$.

Hence we see that the inclusion $E(u)^{-m}\mathfrak{p}/\mathfrak{q} \hookrightarrow (D \otimes_{R\otimes K_0} \mathbb{B}_R)/\mathfrak{q}$ realizes $E(u)^{-m}\mathfrak{p}/\mathfrak{q}$ as a direct summand of the $R$-module $E(u)^{-m}\mathfrak{p}/E(u)^n\mathfrak{p}$ which is locally free by Lemma 2.3(a). This shows that $E(u)^{-m}\mathfrak{p}/E(u)^n\mathfrak{p} \cong (E(u)^{-m}\mathfrak{p}/\mathfrak{q}) \oplus (\mathfrak{q}/E(u)^n\mathfrak{p})$ and both $E(u)^{-m}\mathfrak{p}/\mathfrak{q}$ and $\mathfrak{q}/E(u)^n\mathfrak{p}$ are finite locally free $R$-modules.

Conversely any isomorphism $E(u)^{-m}\mathfrak{p}/E(u)^n\mathfrak{p} \cong (E(u)^{-m}\mathfrak{p}/\mathfrak{q}) \oplus (\mathfrak{q}/E(u)^n\mathfrak{p})$ together with the decomposition $D \otimes_{R\otimes K_0} \mathbb{B}_R \cong (E(u)^n\mathfrak{p}) \oplus (E(u)^{-m}\mathfrak{p}/E(u)^n\mathfrak{p}) \oplus (D \otimes_{R\otimes K_0} \mathbb{B}_R)/E(u)^{-m}\mathfrak{p}$ realizes $\mathfrak{q}$ as a direct summand of $D \otimes_{R\otimes K_0} \mathbb{B}_R$. Indeed, we have the following direct sum decompositions of $R$-modules:

$$\begin{aligned} \mathfrak{q} &\cong (E(u)^n\mathfrak{p}) \oplus (\mathfrak{q}/E(u)^n\mathfrak{p}), \\ D \otimes_{R\otimes K_0} \mathbb{B}_R &\cong (E(u)^n\mathfrak{p}) \oplus (\mathfrak{q}/E(u)^n\mathfrak{p}) \oplus (E(u)^{-m}\mathfrak{p}/\mathfrak{q}) \oplus (D \otimes_{R\otimes K_0} \mathbb{B}_R)/E(u)^{-m}\mathfrak{p}. \end{aligned}$$

Since $E(u)^n\mathfrak{p}$ is finitely generated over $\mathbb{B}^+_R$ and $\mathfrak{q}/E(u)^n\mathfrak{p}$ is finitely generated over $R$, also $\mathfrak{q}$ is finitely generated over $\mathbb{B}^+_R$, hence a $\mathbb{B}^+_R$-lattice.

To prove the local freeness of $\mathfrak{q}$ we may work locally on $R$ and assume by Lemma 2.3(a) that $\mathfrak{p}$ is free over $\mathbb{B}^+_R$, say of rank $d$, and $\mathfrak{q}/E(u)^n\mathfrak{p}$ and $E(u)^{-m}\mathfrak{p}/\mathfrak{q}$ are free over $R$. There is a noetherian subring $\widetilde{R}$ of $R$ and a short exact sequence

$$0 \to \widetilde{Q} \to \widetilde{P} \to \widetilde{N} \to 0 \tag{2-2}$$

of $\mathbb{B}^+_{\widetilde{R}}$-modules which are free $\widetilde{R}$-modules, such that the tensor product of (2-2) with $R$ over $\widetilde{R}$ is isomorphic to

$$0 \to \mathfrak{q}/E(u)^n\mathfrak{p} \to E(u)^{-m}\mathfrak{p}/E(u)^n\mathfrak{p} \to E(u)^{-m}\mathfrak{p}/\mathfrak{q} \to 0. \tag{2-3}$$

Indeed, we can take $\widetilde{R}$ as the finitely generated $\mathbb{Q}_p$-algebra containing all the coefficients appearing in matrix representations of the maps in (2-3) and the action of $K_0[u]/(E(u))^{m+n}$. Note that, since $\widetilde{P}$ is free over $\widetilde{R}$, it is contained in $\widetilde{P} \otimes_{\widetilde{R}} R \cong E(u)^{-m}\mathfrak{p}/E(u)^n\mathfrak{p}$, and since the latter is annihilated by $E(u)^{m+n}$, the same is true for $\widetilde{P}$, $\widetilde{Q}$ and $\widetilde{N}$. Let $\tilde{\mathfrak{p}}$ be a free $\mathbb{B}^+_{\widetilde{R}}$-module of rank $d$ and fix an isomorphism $\tilde{\mathfrak{p}} \otimes_{\mathbb{B}^+_{\widetilde{R}}} \mathbb{B}^+_R \cong \mathfrak{p}$. This isomorphism obviously induces an isomorphism

$$E(u)^{-m}\tilde{\mathfrak{p}} \otimes_{\mathbb{B}^+_{\widetilde{R}}} \left(\mathbb{B}^+_{\widetilde{R}}/(E(u))^{m+n}\right) \xrightarrow{\sim} \widetilde{P}.$$

Let the $\mathbb{B}^+_{\widetilde{R}}$-module $\tilde{\mathfrak{q}}$ be defined by the exact sequence

$$0 \to \tilde{\mathfrak{q}} \to E(u)^{-m}\tilde{\mathfrak{p}} \to \widetilde{N} \to 0. \tag{2-4}$$

Since $\mathbb{B}^+_{\widetilde{R}} \cong (\widetilde{R} \otimes_{\mathbb{Q}_p} K)[\![t]\!]$ is noetherian, $\tilde{\mathfrak{q}}$ is finitely generated. Consider a maximal ideal $\mathfrak{m} \subset \mathbb{B}^+_{\widetilde{R}}$. Since $t \in \mathfrak{m}$, it maps to a maximal ideal $\mathfrak{n}$ of $\widetilde{R}$. Since $\mathfrak{n}$ is finitely generated, $\mathbb{B}^+_{\widetilde{R}} \otimes_{\widetilde{R}} \widetilde{R}/\mathfrak{n} \cong (\widetilde{R}/\mathfrak{n} \otimes_{\mathbb{Q}_p} K)[\![t]\!]$ and this is a direct product of discrete valuation rings. Thus $\tilde{\mathfrak{q}} \otimes_{\widetilde{R}} \widetilde{R}/\mathfrak{n}$ is locally free of rank $d$ by the elementary divisor theorem. Since this holds for all $\mathfrak{m}$, [EGA IV$_3$ 1966, Theorem 11.3.10] implies that $\tilde{\mathfrak{q}}$ is a projective $\mathbb{B}^+_{\widetilde{R}}$-module and by [EGA I 1971, Proposition 10.10.8.6] it is locally on Spec $\widetilde{R} \otimes_{\mathbb{Q}_p} K$ free over $\mathbb{B}^+_{\widetilde{R}}$. Let $\{\psi : K \hookrightarrow \overline{\mathbb{Q}}_p\}$ be the set of all $\mathbb{Q}_p$-homomorphisms and let $\widetilde{K}$ be the compositum of all $\psi(K)$ inside $\overline{\mathbb{Q}}_p$. Then $\widetilde{R} \to \widetilde{R} \otimes_{\mathbb{Q}_p} \widetilde{K}$ is finite étale and the pullback of $\tilde{\mathfrak{q}}$ under this base change is locally on Spec $\widetilde{R} \otimes_{\mathbb{Q}_p} \widetilde{K} \otimes_{\mathbb{Q}_p} K$ free over $\mathbb{B}^+_{\widetilde{R} \otimes_{\mathbb{Q}_p} \widetilde{K}}$. Since

$$\operatorname{Spec} \widetilde{R} \otimes_{\mathbb{Q}_p} \widetilde{K} \otimes_{\mathbb{Q}_p} K = \coprod_{\psi} \operatorname{Spec} \widetilde{R} \otimes_{\mathbb{Q}_p} \widetilde{K},$$

the pullback of $\tilde{\mathfrak{q}}$ is already locally on Spec $\widetilde{R} \otimes_{\mathbb{Q}_p} \widetilde{K}$ free over $\mathbb{B}^+_{\widetilde{R} \otimes_{\mathbb{Q}_p} \widetilde{K}}$.

To finish the proof it remains to show that $\tilde{\mathfrak{q}} \otimes_{\mathbb{B}^+_{\widetilde{R}}} \mathbb{B}^+_R \cong \mathfrak{q}$. Tensoring (2-4) with $\mathbb{B}^+_R$ over $\mathbb{B}^+_{\widetilde{R}}$ we obtain the top row in the diagram

$$\begin{array}{ccccccccc}
0 & \longrightarrow & \operatorname{Tor}_1^{\mathbb{B}^+_{\widetilde{R}}}(\widetilde{N}, \mathbb{B}^+_R) & \longrightarrow & \tilde{\mathfrak{q}} \otimes_{\mathbb{B}^+_{\widetilde{R}}} \mathbb{B}^+_R & \longrightarrow & E(u)^{-m}\mathfrak{p} & \longrightarrow & \widetilde{N} \otimes_{\mathbb{B}^+_{\widetilde{R}}} \mathbb{B}^+_R \longrightarrow 0 \\
 & & & & \downarrow & & \| & & \downarrow \cong \\
0 & \longrightarrow & & & \mathfrak{q} & \longrightarrow & E(u)^{-m}\mathfrak{p} & \longrightarrow & E(u)^{-m}\mathfrak{p}/\mathfrak{q} \longrightarrow 0
\end{array}$$

Abbreviate $\ell := m + n$. Since the functor $\widetilde{N} \otimes_{\mathbb{B}^+_{\widetilde{R}}} \bullet$ equals the composition of the functors $(\mathbb{B}^+_{\widetilde{R}}/t^\ell) \otimes_{\mathbb{B}^+_{\widetilde{R}}} \bullet$ followed by $\widetilde{N} \otimes_{\mathbb{B}^+_{\widetilde{R}}/t^\ell} \bullet$, the Tor$_1$-module on the left can be computed from a change of rings spectral sequence [Rotman 2009, Theorem 10.71] and its associated 5-term sequence of low degrees, see [Rotman 2009, Theorem 10.31],

$$\cdots \to \operatorname{Tor}_1^{\mathbb{B}^+_{\widetilde{R}}}(\mathbb{B}^+_{\widetilde{R}}/t^\ell, \mathbb{B}^+_R) \otimes_{\mathbb{B}^+_{\widetilde{R}}/t^\ell} \widetilde{N} \to \operatorname{Tor}_1^{\mathbb{B}^+_{\widetilde{R}}}(\widetilde{N}, \mathbb{B}^+_R) \to \operatorname{Tor}_1^{\mathbb{B}^+_{\widetilde{R}}/t^\ell}(\widetilde{N}, \mathbb{B}^+_R/t^\ell) \to 0.$$

The right term in this sequence is zero because $\mathrm{Tor}_1^{\mathbb{B}_{\widetilde{R}}^+/t^\ell}(\widetilde{N}, \mathbb{B}_R^+/t^\ell) = \mathrm{Tor}_1^{\widetilde{R}}(\widetilde{N}, R)$ and $\widetilde{N}$ is flat over $\widetilde{R}$. The left term is zero because $t^\ell$ is a nonzero-divisor both in $\mathbb{B}_{\widetilde{R}}^+$ and $\mathbb{B}_R^+$. This shows that $\mathrm{Tor}_1^{\mathbb{B}_{\widetilde{R}}^+}(\widetilde{N}, \mathbb{B}_R^+) = 0$ and proves the lemma. □

**Remark 2.8.** (1) Let $R = L$ be a field and let $(D, \Phi, N, \mathfrak{q})$ be a $(\varphi, N)$-module with Hodge–Pink lattice over $L$. The Hodge–Pink lattice $\mathfrak{q}$ gives rise to a $K$-filtration $\mathcal{F}_{\mathfrak{q}}^\bullet$ as follows. Consider the natural projection

$$\begin{aligned}\mathfrak{p} \twoheadrightarrow \mathfrak{p}/E(u)\mathfrak{p} &= D \otimes_{R\otimes K_0} \mathbb{B}_R^+/(E(u)) \\ &= D \otimes_{R\otimes K_0} R \otimes_{\mathbb{Q}_p} K = D_K\end{aligned}$$

and let $\mathcal{F}_{\mathfrak{q}}^i D_K$ be the image of $\mathfrak{p} \cap E(u)^i\mathfrak{q}$ in $D_K$ for all $i \in \mathbb{Z}$, that is

$$\mathcal{F}_{\mathfrak{q}}^i D_K := \big(\mathfrak{p} \cap E(u)^i\mathfrak{q}\big)\big/\big(E(u)\mathfrak{p} \cap E(u)^i\mathfrak{q}\big).$$

Since $L$ is a field $(D, \Phi, N, \mathcal{F}_{\mathfrak{q}}^\bullet)$ is a $K$-filtered $(\varphi, N)$-module over $L$. Note that this functor does not exist for general $R$, because $\mathrm{gr}^i_{\mathcal{F}_{\mathfrak{q}}} D_K$ will not be locally free over $R$ in general. This is related to the fact that the Hodge polygon of $\mathcal{F}_{\mathfrak{q}}^\bullet$ is locally constant on $R$ whereas the Hodge polygon of $\mathfrak{q}$ is only semicontinuous; see Remark 2.12 below.

(2) However, for general $R$ consider the category of $(\varphi, N)$-modules with Hodge–Pink lattice $(D, \Phi, N, \mathfrak{q})$ over $R$, such that $\mathfrak{p} \subset \mathfrak{q} \subset E(u)^{-1}\mathfrak{p}$. This category is equivalent to the category of $K$-filtered $(\varphi, N)$-modules $(D, \Phi, N, \mathcal{F}^\bullet)$ over $R$ with $\mathcal{F}^0 D_K = D_K$ and $\mathcal{F}^2 = 0$. Namely, defining $\mathcal{F}_{\mathfrak{q}}^\bullet$ as in (1) we obtain

$$\mathrm{gr}^i_{\mathcal{F}_{\mathfrak{q}}} D_K \cong \begin{cases} E(u)^{-1}\mathfrak{p}/\mathfrak{q} & \text{for } i = 0, \\ \mathfrak{q}/\mathfrak{p} & \text{for } i = 1, \\ 0 & \text{for } i \neq 0, 1, \end{cases}$$

and so $(D, \Phi, N, \mathcal{F}_{\mathfrak{q}}^\bullet)$ is a $K$-filtered $(\varphi, N)$-module by Lemma 2.7. Conversely, $\mathfrak{q}$ equals the preimage of $\mathcal{F}_{\mathfrak{q}}^1 D_K$ under the morphism $E(u)^{-1}\mathfrak{p} \xrightarrow{\cdot E(u)} \mathfrak{p} \twoheadrightarrow D_K$ and this defines the inverse functor.

(3) Now let $(D, \Phi, N, \mathcal{F}^\bullet)$ be a $K$-filtered $(\varphi, N)$-module over $R$. Using that $\mathbb{B}_R^+ = (R \otimes_{\mathbb{Q}_p} K)[\![t]\!]$ is an $R \otimes_{\mathbb{Q}_p} K$-algebra, we can define the Hodge–Pink lattice

$$\mathfrak{q} := \mathfrak{q}(\mathcal{F}^\bullet) := \sum_{i\in\mathbb{Z}} E(u)^{-i}(\mathcal{F}^i D_K) \otimes_{R\otimes K} \mathbb{B}_R^+.$$

It satisfies $\mathcal{F}_{\mathfrak{q}}^\bullet = \mathcal{F}^\bullet$. Using Lemma 2.7 one easily finds that $\mathfrak{q}(\mathcal{F}^\bullet)$ is indeed a $\mathbb{B}_R^+$-lattice.

**Example 2.9.** The $K$-filtered $(\varphi, N)$-modules over $R = \mathbb{Q}_p$ which correspond to the cyclotomic character $\chi_{\mathrm{cyc}} : \mathscr{G}_K \to \mathbb{Z}_p^\times$ are $D_{\mathrm{st}}(\chi_{\mathrm{cyc}}) = (K_0, \Phi = p^{-1}, N = 0, \mathcal{F}^\bullet)$ with $\mathcal{F}^{-1} = K \supsetneq \mathcal{F}^0 = (0)$ and its dual $D_{\mathrm{st}}^*(\chi_{\mathrm{cyc}}) = (K_0, \Phi = p, N = 0, \mathcal{F}^\bullet)$ with $\mathcal{F}^1 = K \supsetneq \mathcal{F}^2 = (0)$. For both there exists a unique Hodge–Pink lattice which induces the filtration. On $D_{\mathrm{st}}(\chi_{\mathrm{cyc}})$ it is $\mathfrak{q} = E(u)\mathfrak{p}$ and on $D_{\mathrm{st}}^*(\chi_{\mathrm{cyc}})$ it is $\mathfrak{q} = E(u)^{-1}\mathfrak{p}$.

We want to introduce Hodge weights and Hodge polygons. Let $d > 0$, let $B \subset \mathrm{GL}_d$ be the Borel subgroup of upper triangular matrices and let $T \subset B$ be the maximal torus consisting of the diagonal

matrices. Let $\widetilde{G} := \mathrm{Res}_{K/\mathbb{Q}_p} \mathrm{GL}_{d,K}$, $\widetilde{B} = \mathrm{Res}_{K/\mathbb{Q}_p} B$ and $\widetilde{T} := \mathrm{Res}_{K/\mathbb{Q}_p} T$ be the Weil restrictions. We consider cocharacters

$$\mu : \mathbb{G}_{m,\overline{\mathbb{Q}}_p} \to \widetilde{T}_{\overline{\mathbb{Q}}_p} \tag{2-5}$$

which are dominant with respect to the Borel $\widetilde{B}$ of $\widetilde{G}$. In other words on $\overline{\mathbb{Q}}_p$-valued points the cocharacter

$$\mu : \overline{\mathbb{Q}}_p^\times \to \prod_{\psi : K \to \overline{\mathbb{Q}}_p} T(\overline{\mathbb{Q}}_p)\,,$$

where $\psi$ runs over all $\mathbb{Q}_p$-homomorphisms $\psi : K \to \overline{\mathbb{Q}}_p$, is given by cocharacters

$$\mu_\psi : x \mapsto \mathrm{diag}(x^{\mu_{\psi,1}}, \ldots, x^{\mu_{\psi,d}})$$

for some integers $\mu_{\psi,j} \in \mathbb{Z}$ with $\mu_{\psi,j} \geq \mu_{\psi,j+1}$. We define the *reflex field* $E_\mu$ *of* $\mu$ as the fixed field in $\overline{\mathbb{Q}}_p$ of $\{\sigma \in \mathscr{G}_{\mathbb{Q}_p} : \mu_{\sigma\psi,j} = \mu_{\psi,j} \text{ for all } j, \psi\}$. It is a finite extension of $\mathbb{Q}_p$ which is contained in the compositum $\widetilde{K}$ of all $\psi(K)$ inside $\overline{\mathbb{Q}}_p$. For each $j$ the locally constant function $\psi \mapsto \mu_{\psi,j}$ on $\mathrm{Spec}\, \widetilde{K} \otimes_{\mathbb{Q}_p} K \cong \coprod_{\psi : K \to \widetilde{K}} \mathrm{Spec}\, \widetilde{K}$ descends to a $\mathbb{Z}$-valued function $\mu_j$ on $\mathrm{Spec}\, E_\mu \otimes_{\mathbb{Q}_p} K$, because $\mu_j$ is constant on the fibers of $\mathrm{Spec}\, \widetilde{K} \otimes_{\mathbb{Q}_p} K \to \mathrm{Spec}\, E_\mu \otimes_{\mathbb{Q}_p} K$. In particular, the cocharacter $\mu$ is defined over $E_\mu$. If $R$ is an $E_\mu$-algebra we also view $\mu_j$ as a locally constant $\mathbb{Z}$-valued function on $\mathrm{Spec}\, R \otimes_{\mathbb{Q}_p} K$.

**Construction 2.10.** Let $\underline{D} = (D, \Phi, N, \mathfrak{q})$ be a $(\varphi, N)$-module with Hodge–Pink lattice of rank $d$ over a field extension $L$ of $\mathbb{Q}_p$. By Lemma 2.3(a) the $L \otimes_{\mathbb{Q}_p} K_0$-module $D$ is free. Since $L \otimes_{\mathbb{Q}_p} K$ is a product of fields, $\mathbb{B}_L^+ = (L \otimes_{\mathbb{Q}_p} K)[\![t]\!]$ is a product of discrete valuation rings and $\mathfrak{q}$ is a free $\mathbb{B}_L^+$-module of rank $d$. We choose bases of $D$ and $\mathfrak{q}$. Then the inclusion $\mathfrak{q} \subset D \otimes_{L \otimes K_0} \mathbb{B}_L$ is given by an element $\gamma$ of $\mathrm{GL}_d(\mathbb{B}_L) = \widetilde{G}\big(L(\!(t)\!)\big)$. By the Cartan decomposition for $\widetilde{G}$ there is a uniquely determined dominant cocharacter $\mu_L : \mathbb{G}_{m,L} \to \widetilde{T}_L$ over $L$ with $\gamma \in \widetilde{G}\big(L[\![t]\!]\big)\mu_L(t)^{-1}\widetilde{G}\big(L[\![t]\!]\big)$. This cocharacter is independent of the chosen bases. If $L$ contains $\widetilde{K}$, it is defined over $\widetilde{K}$ because $\widetilde{T}$ splits over $\widetilde{K}$. In this case we view it as an element of $X_*(T_{\widetilde{K}})_{\mathrm{dom}}$ and denote it by $\mu_{\underline{D}}(\mathrm{Spec}\, L)$. It has the following explicit description. Under the decomposition $L \otimes_{\mathbb{Q}_p} K = \prod_{\psi : K \to \widetilde{K}} L$ we have $\gamma \in \prod_\psi \mathrm{GL}_d\big(L[\![t]\!]\big)\mu_\psi(t)^{-1}\,\mathrm{GL}_d\big(L[\![t]\!]\big)$, as $\widetilde{G}\big(L(\!(t)\!)\big) = \prod_\psi \mathrm{GL}_d\big(L(\!(t)\!)\big)$, and $\mu_L = (\mu_\psi)_\psi$. The $t^{-\mu_{\psi,1}}, \ldots, t^{-\mu_{\psi,d}}$ are the elementary divisors of the $\psi$-component $\mathfrak{q}_\psi$ of $\mathfrak{q}$ with respect to $\mathfrak{p}$. That is, there is an $L[\![t]\!]$-basis $(v_{\psi,1}, \ldots, v_{\psi,d})$ of the $\psi$-component $\mathfrak{p}_\psi$ of $\mathfrak{p}$ such that $(t^{-\mu_{\psi,1}} v_{\psi,1}, \ldots, t^{-\mu_{\psi,d}} v_{\psi,d})$ is an $L[\![t]\!]$-basis of $\mathfrak{q}_\psi$.

Let $(D, \Phi, N, \mathcal{F}^\bullet_{\mathfrak{q}})$ be the $K$-filtered $(\varphi, N)$-module associated with $\underline{D}$ by Remark 2.8(1). Then $\mathcal{F}^i_{\mathfrak{q}} D_{K,\psi} = \langle v_{\psi,j} : i - \mu_{\psi,j} \leq 0\rangle_L$ and

$$\dim_L \mathrm{gr}^i_{\mathcal{F}_{\mathfrak{q}}} D_{K,\psi} = \#\{j : i - \mu_{\psi,j} = 0\}.$$

More generally, for a $K$-filtered $(\varphi, N)$-module $(D, \Phi, N, \mathcal{F}^\bullet)$ over a field extension $L$ of $\widetilde{K}$ we consider the decomposition $D_K = \prod_\psi D_{K,\psi}$ and define the integers $\mu_{\psi,1} \geq \cdots \geq \mu_{\psi,d}$ by the formula

$$\dim_L \mathrm{gr}^i_{\mathcal{F}} D_{K,\psi} = \#\{j : \mu_{\psi,j} = i\}.$$

We define the cocharacter $\mu_{(D,\Phi,N,\mathcal{F}^\bullet)}(\mathrm{Spec}\, L) := (\mu_\psi)_\psi$ and view it as an element of $X_*(T_{\widetilde{K}})_{\mathrm{dom}}$.

**Definition 2.11.** (a) Let $R$ be a $\widetilde{K}$-algebra and consider the decomposition $R \otimes_{\mathbb{Q}_p} K = \prod_{\psi: K \hookrightarrow \widetilde{K}} R$. Let $\underline{D}$ be a $(\varphi, N)$-module with Hodge–Pink lattice (respectively a $K$-filtered $(\varphi, N)$-module) of rank $d$ over $R$. For every point $s \in \operatorname{Spec} R$ we consider the base change $s^*\underline{D}$ of $\underline{D}$ to $\kappa(s)$. We call the cocharacter $\mu_{\underline{D}}(s) := \mu_{s^*\underline{D}}(\operatorname{Spec} \kappa(s))$ from Construction 2.10 the *Hodge polygon of* $\underline{D}$ *at* $s$ and we consider $\mu_{\underline{D}}$ as a function $\mu_{\underline{D}} : \operatorname{Spec} R \to X_*(T_{\widetilde{K}})_{\mathrm{dom}}$. The integers $-\mu_{\psi,j}(s)$ are called the *Hodge weights of* $\underline{D}$ *at* $s$.

Now let $\mu : \overline{\mathbb{Q}}_p^\times \to \widetilde{T}(\overline{\mathbb{Q}}_p)$ be a dominant cocharacter as in (2-5), let $E_\mu$ denote the reflex field of $\mu$, and let $R$ be an $E_\mu$-algebra.

(b) Let $\underline{D}$ be a $(\varphi, N)$-module with Hodge–Pink lattice (respectively a $K$-filtered $(\varphi, N)$-module) of rank $d$ over $R$. We say that $\underline{D}$ *has constant Hodge polygon equal to* $\mu$ if $\mu_{\underline{D}}(s) = \mu$ for every point $s \in \operatorname{Spec}(R \otimes_{E_\mu} \widetilde{K})$.

(c) Let $\underline{D} = (D, \Phi, N, \mathfrak{q})$ be a $(\varphi, N)$-module with Hodge–Pink lattice over $\operatorname{Spec} R$. We say that $\underline{D}$ has *Hodge polygon bounded by* $\mu$ if

$$\textstyle\bigwedge^j_{\mathbb{B}^+_R} \mathfrak{q} \subset E(u)^{-\mu_1 - \cdots - \mu_j} \cdot \bigwedge^j_{\mathbb{B}^+_R} \mathfrak{p}$$

for all $j = 1, \ldots, d$ with equality for $j = d$, where the $\mu_i$ are the $\mathbb{Z}$-valued functions on $\operatorname{Spec} R \otimes_{\mathbb{Q}_p} K$ determined by $\mu$; see the discussion before Construction 2.10.

Equivalently the condition of being bounded by $\mu$ can be described as follows: Over $\widetilde{K}$ the cocharacter $\mu$ is described by a decreasing sequence of integers $\mu_{\psi,1} \geq \cdots \geq \mu_{\psi,d}$ for every $\mathbb{Q}_p$-embedding $\psi : K \hookrightarrow \overline{\mathbb{Q}}_p$. Let $R' = R \otimes_{E_\mu} \widetilde{K}$, then $R' \otimes_{\mathbb{Q}_p} K \cong \prod_{\psi: K \to \widetilde{K}} R'_\psi$ with each $R'_\psi = R'$ under the isomorphism $a \otimes b \mapsto (a\psi(b))_\psi$, where $\psi : K \to R'$ is given via the embedding into the second factor of $R' = R \otimes_{E_\mu} \widetilde{K}$. Especially we view $R'_\psi$ as a $K$-algebra via $\psi$. Under this isomorphism $D \otimes_{R \otimes K_0} \mathbb{B}_{R'} =: \mathfrak{p}_{R'}[1/t]$ decomposes into a product $\prod_\psi \mathfrak{p}_{R'}[1/t]_\psi$, where $\mathfrak{p}_{R'}[1/t]_\psi$ is a free $R'_\psi[\![t]\!][1/t]$-module and the $\mathbb{B}^+_{R'}$-lattice $\mathfrak{p}_{R'} \subset \mathfrak{p}_{R'}[1/t]$ decomposes into a product of $R'_\psi[\![t]\!]$-lattices $\mathfrak{p}_{R',\psi} \subset \mathfrak{p}_{R'}[1/t]_\psi$.

Further, under the isomorphism $D \otimes_{R \otimes K_0} \mathbb{B}_{R'} \cong \prod_\psi \mathfrak{p}_{R'}[1/t]_\psi$ the Hodge–Pink lattice $\mathfrak{q}_{R'} = \mathfrak{q} \otimes_R R'$ decomposes into a product $\mathfrak{q}_{R'} = \prod_\psi \mathfrak{q}_{R',\psi}$, where $\mathfrak{q}_{R',\psi}$ is an $R'_\psi[\![t]\!]$-lattice in $\mathfrak{p}_{R'}[1/t]_\psi$. Then the condition of being bounded by $\mu$ is equivalent to

$$\textstyle\bigwedge^j_{\mathbb{B}^+_{R'}} \mathfrak{q}_{R',\psi} \subset E(u)^{-\mu_{\psi,1} - \cdots - \mu_{\psi,j}} \cdot \bigwedge^j_{\mathbb{B}^+_{R'}} \mathfrak{p}_{R',\psi} \tag{2-6}$$

for all $\psi$ and all $j = 1, \ldots, d$ with equality for $j = d$.

Note that by Cramer's rule (e.g., [Bourbaki 1970, III.8.6, Formulas (21) and (22)]) the condition of Definition 2.11 (c), respectively (2-6) is equivalent to

$$\textstyle\bigwedge^j_{\mathbb{B}^+_R} \mathfrak{p} \subset E(u)^{\mu_{d-j+1} + \cdots + \mu_d} \cdot \bigwedge^j_{\mathbb{B}^+_R} \mathfrak{q}\,,$$

respectively

$$\textstyle\bigwedge^j_{\mathbb{B}^+_{R'}} \mathfrak{p}_{R',\psi} \subset E(u)^{\mu_{\psi,d-j+1} + \cdots + \mu_{\psi,d}} \cdot \bigwedge^j_{\mathbb{B}^+_{R'}} \mathfrak{q}_{R',\psi} \tag{2-7}$$

for all $j = 1, \ldots, d$ with equality for $j = d$.

**Remark 2.12.** The Hodge polygon of a $K$-filtered $(\varphi, N)$-module $(D, \Phi, N, \mathcal{F}^\bullet)$ is locally constant on $R$, because $\mathrm{gr}^i_{\mathcal{F}} D_{K,\psi}$ is locally free over $R$ as a direct summand of the locally free $R$-module $\mathrm{gr}^i_{\mathcal{F}} D_K$.

In contrast, the Hodge polygon of a $(\varphi, N)$-module $\underline{D}$ with Hodge–Pink lattice over $R$ is not locally constant in general. Nevertheless, for any cocharacter $\mu$ as in (2-5) the set of points $s \in \operatorname{Spec} R$ such that $\mu_{\underline{D}}(s) \preceq \mu$ in the Bruhat order, is closed in $\operatorname{Spec} R$. This is a consequence of the next:

**Proposition 2.13.** *Let $\mu \in X_*(T_{\widetilde{K}})_{\mathrm{dom}}$ be a dominant cocharacter with reflex field $E_\mu$ and let $R$ be an $E_\mu$-algebra. Let $\underline{D} = (D, \Phi, N, \mathfrak{q})$ be a $(\varphi, N)$-module with Hodge–Pink lattice of rank $d$ over $R$.*

(a) *The condition that $\underline{D}$ has Hodge polygon bounded by $\mu$ is representable by a finitely presented closed immersion* $(\operatorname{Spec} R)_{\preceq\mu} \hookrightarrow \operatorname{Spec} R$.

(b) *If $R$ is reduced then $\underline{D}$ has Hodge polygon bounded by $\mu$ if and only if for all points $s \in \operatorname{Spec} R \otimes_{E_\mu} \widetilde{K}$ we have $\mu_{\underline{D}}(s) \preceq \mu$ in the Bruhat order, that is, for all $\psi$ the vector $\mu_\psi - \mu_{\underline{D}}(s)_\psi \in \mathbb{Z}^d$ is a nonnegative linear combination of the positive coroots $\check{\alpha}_j = (\ldots, 0, 1, -1, 0, \ldots)$ having the "1" as the $j$-th entry.*

(c) *Let $\mu'$ be another dominant cocharacter such that $\mu' \preceq \mu$ in the Bruhat order. Let $E_{\mu'}$ denote its reflex field and let $E = E_\mu E_{\mu'} \subset \widetilde{K}$ be the composite field. Assume that $R$ is an $E$-algebra, then* $(\operatorname{Spec} R)_{\preceq\mu'} \hookrightarrow (\operatorname{Spec} R)_{\preceq\mu}$ *as closed subschemes of* $\operatorname{Spec} R$.

*Proof.* (a) By Lemma 2.7 we find a large positive integer $n$ such that $E(u)^n\mathfrak{p} \subset \mathfrak{q} \subset E(u)^{-n}\mathfrak{p}$. This implies $\bigwedge^j \mathfrak{q} \subset E(u)^{-jn} \bigwedge^j \mathfrak{p}$ for all $j$ and $\bigwedge^d \mathfrak{p} \subset E(u)^{-dn} \bigwedge^d \mathfrak{q}$. Viewing $\mu_j : \operatorname{Spec} R \otimes_{\mathbb{Q}_p} K \to \mathbb{Z}$ as locally constant function as in the discussion before Construction 2.10, we consider the modules over $\mathbb{B}^+_R \cong (R \otimes_{\mathbb{Q}_p} K)[\![t]\!]$

$$\begin{aligned} M_0 &:= E(u)^{-dn} \textstyle\bigwedge^d \mathfrak{q} \,\big/\, E(u)^{\mu_1+\cdots+\mu_d} \cdot \bigwedge^d \mathfrak{q}, \\ M_j &:= E(u)^{-jn} \textstyle\bigwedge^j \mathfrak{p} \,\big/\, E(u)^{-\mu_1-\cdots-\mu_j} \cdot \bigwedge^j \mathfrak{p} \qquad \text{for } 1 \le j \le d. \end{aligned} \tag{2-8}$$

As $R$-modules they are finite locally free. Then $\underline{D}$ has Hodge polygon bounded by $\mu$ if and only if for all $j = 1, \ldots, d$ all generators of $\bigwedge^j \mathfrak{q}$ are mapped to zero in $M_j$ and all generators of $\bigwedge^d \mathfrak{p}$ are mapped to zero in $M_0$. Since $M := M_0 \oplus \cdots \oplus M_d$ is finite locally free over $R$, this condition is represented by a finitely presented closed immersion into $\operatorname{Spec} R$ by [EGA I 1971, Lemma 9.7.9.1].

(b) If $R$ is reduced then also the étale $R$-algebra $R' := R \otimes_{E_\mu} \widetilde{K}$ is reduced and $R \hookrightarrow R' \hookrightarrow \prod_{s \in \operatorname{Spec} R'} \kappa(s)$ is injective. Therefore also $M \hookrightarrow M \otimes_R \big(\prod_{s \in \operatorname{Spec} R'} \kappa(s)\big)$ is injective. So $\underline{D}$ has Hodge polygon bounded by $\mu$ if and only if this holds for the pullbacks $s^*\underline{D}$ to $\operatorname{Spec} \kappa(s)$ at all points $s \in \operatorname{Spec} R'$. By definition of $\mu' := \mu_{\underline{D}}(s)$ there is a $\kappa(s)[\![t]\!]$-basis $(v_{\psi,1}, \ldots, v_{\psi,d})$ of the $\psi$-component $(s^*\mathfrak{p})_\psi$ of $s^*\mathfrak{p}$ such that $(t^{-\mu'_{\psi,1}} v_{\psi,1}, \ldots, t^{-\mu'_{\psi,d}} v_{\psi,d})$ is a $\kappa(s)[\![t]\!]$-basis of $(s^*\mathfrak{q})_\psi$. Therefore condition (2-6) holds if and only if $\mu_{\psi,1} + \cdots + \mu_{\psi,j} \ge \mu'_{\psi,1} + \cdots + \mu'_{\psi,j}$ for all $\psi$ and $j$ with equality for $j = d$. One easily checks that this is equivalent to $\mu' \preceq \mu$.

(c) Again $\mu' \preceq \mu$ implies $\mu_{\psi,1} + \cdots + \mu_{\psi,j} \ge \mu'_{\psi,1} + \cdots + \mu'_{\psi,j}$ for all $\psi$ and $j$ with equality for $j = d$. We view $\mu_j, \mu'_j$ as locally constant $\mathbb{Z}$-valued functions on $\operatorname{Spec} E \otimes_{\mathbb{Q}_p} K$. Then $\mu_1 + \cdots + \mu_j \ge \mu'_1 + \cdots + \mu'_j$

for all $j$ with equality for $j = d$. In terms of (2-8) the $R$-modules $M_j$ for $\mu$ are quotients of the $R$-modules $M'_j$ for $\mu'$ with $M'_0 = M_0$. Therefore $(\operatorname{Spec} R)_{\preceq\mu'} \hookrightarrow \operatorname{Spec} R$ factors through $(\operatorname{Spec} R)_{\preceq\mu}$. □

**Remark 2.14.** The reader should note that $\mu' \preceq \mu$ does not imply a relation between $E_{\mu'}$ and $E_\mu$ as can be seen from the following example. Let $d = 2$ and $[K : \mathbb{Q}_p] = 2$ and $\{\psi : K \hookrightarrow \widetilde{K}\} = \operatorname{Gal}(K/\mathbb{Q}_p) = \{\psi_1, \psi_2\}$. Consider the three cocharacters $\mu, \mu', \mu''$ given by $\mu_{\psi_1} = (2,0),\ \mu_{\psi_2} = (2,0)$ and $\mu'_{\psi_1} = (2,0),\ \mu'_{\psi_2} = (1,1)$ and $\mu''_{\psi_1} = (1,1),\ \mu''_{\psi_2} = (1,1)$. Then $\mu'' \preceq \mu' \preceq \mu$. On the other hand we find $E_\mu = E_{\mu''} = \mathbb{Q}_p$ and $E_{\mu'} = \widetilde{K} = K$.

**Remark 2.15.** In Definition 2.11 (a) we assumed that $R$ is a $\widetilde{K}$-algebra to obtain a well defined Hodge polygon $\mu_{\underline{D}}(s) \in X_*(\widetilde{T}_{\widetilde{K}})$. In Definition 2.11 (b) we can lower the ground field over which $R$ is defined to $E_\mu$ because $\operatorname{Gal}(\widetilde{K}/E_\mu)$ fixes $\mu$. The ground field cannot be lowered further, as one sees from the following:

**Proposition 2.16.** *Let $\underline{D}$ be a $(\varphi, N)$-module with Hodge–Pink lattice (or a $K$-filtered $(\varphi, N)$-module) of rank $d$ over a field $L$ such that $\mu_{\underline{D}}(s) = \mu$ for all points $s \in \operatorname{Spec} L \otimes_{\mathbb{Q}_p} \widetilde{K}$. Then there is a canonical inclusion of the reflex field $E_\mu \hookrightarrow L$.*

*Proof.* Since every $K$-filtered $(\varphi, N)$-module arises from a $(\varphi, N)$-module with Hodge–Pink lattice as in Remark 2.8 (3), it suffices to treat the case where $\underline{D}$ is a $(\varphi, N)$-module with Hodge–Pink lattice. We consider the decomposition $\widetilde{L} := L \otimes_{\mathbb{Q}_p} \widetilde{K} = \prod_{s \in \operatorname{Spec} \widetilde{L}} \kappa(s)$ and for each $s$ we denote by $\alpha_s : L \hookrightarrow \kappa(s)$ and $\beta_s : \widetilde{K} \hookrightarrow \kappa(s)$ the induced inclusions. Let $\mu_L : \mathbb{G}_{m,L} \to \widetilde{T}_L$ be the cocharacter over $L$ associated with $\underline{D}$ in Construction 2.10. The assumption of the proposition means that $\alpha_s(\mu_L) = \beta_s(\mu)$ for all $s$. The Galois group $\mathcal{G} := \operatorname{Gal}(\widetilde{K}, \mathbb{Q}_p)$ acts on $\widetilde{L}$. The Galois group $\operatorname{Gal}(\kappa(s)/\alpha_s(L))$ can be identified with the decomposition group $\mathcal{G}_s := \{\sigma \in \mathcal{G} : \sigma(s) = s\}$ under the monomorphism $\operatorname{Gal}(\kappa(s)/\alpha_s(L)) \hookrightarrow \mathcal{G},\ \tau \mapsto \beta_s^{-1} \circ \tau|_{\beta_s(\widetilde{K})} \circ \beta_s$. Since $\mu_L$ is defined over $L$, each $\tau \in \operatorname{Gal}(\kappa(s)/\alpha_s(L))$ satisfies $\tau(\alpha_s(\mu_L)) = \alpha_s(\mu_L)$, and hence $(\beta_s^{-1} \circ \tau|_{\beta_s(\widetilde{K})} \circ \beta_s)(\mu) = \mu$. By definition of the reflex field $E_\mu$ this implies that $\beta_s^{-1} \circ \tau|_{\beta_s(\widetilde{K})} \circ \beta_s \in \operatorname{Gal}(\widetilde{K}/E_\mu)$ and $\tau|_{\beta_s(E_\mu)} = \mathrm{id}$. So $\beta_s(E_\mu) \subset \alpha_s(L)$ and we get an inclusion $\alpha_s^{-1}\beta_s : E_\mu \hookrightarrow L$. To see that this is independent of $s$ choose a $\sigma \in \mathcal{G}$ with $\sigma(s) = \tilde{s}$. Then $\alpha_{\tilde{s}} = \sigma \circ \alpha_s$ and $\beta_{\tilde{s}} = \sigma \circ \beta_s$. □

## 3. Moduli spaces for $(\varphi, N)$-modules with Hodge–Pink lattice

We will introduce and study moduli spaces for the objects introduced in Section 2. Proposition 2.16 suggests to work over the reflex field.

**Definition 3.1.** Let $\mu$ be a cocharacter as in (2-5) and let $E_\mu$ be its reflex field. We define fpqc-stacks $\mathscr{D}_{\varphi,N,\mu}$, $\mathscr{H}_{\varphi,N,\preceq\mu}$, and $\mathscr{H}_{\varphi,N,\mu}$ on the category of $E_\mu$-schemes. For an affine $E_\mu$-scheme $\operatorname{Spec} R$:

(a) The groupoid $\mathscr{D}_{\varphi,N,\mu}(\operatorname{Spec} R)$ consists of $K$-filtered $(\varphi,N)$-modules $(D, \Phi, N, \mathcal{F}^\bullet)$ over $R$ of rank $d$ with constant Hodge polygon equal to $\mu$.

(b) The groupoid $\mathscr{H}_{\varphi,N,\preceq\mu}(\operatorname{Spec} R)$ consists of $(\varphi, N)$-modules with Hodge–Pink lattice $(D, \Phi, N, \mathfrak{q})$ over $R$ of rank $d$ with Hodge polygon bounded by $\mu$.

(c) The groupoid $\mathscr{H}_{\varphi,N,\mu}(\operatorname{Spec} R)$ consists of $(\varphi, N)$-modules with Hodge–Pink lattice $(D, \Phi, N, \mathfrak{q})$ over $R$ of rank $d$ with Hodge polygon bounded by $\mu$ and constant equal to $\mu$.

Let $\mathscr{D}_{\varphi,\mu} \subset \mathscr{D}_{\varphi,N,\mu}$ (resp. $\mathscr{H}_{\varphi,\preceq\mu} \subset \mathscr{H}_{\varphi,N,\preceq\mu}$, $\mathscr{H}_{\varphi,\mu} \subset \mathscr{H}_{\varphi,N,\mu}$) be the closed substack on which $N$ is zero. It classifies $\varphi$-modules with $K$-filtration (resp. Hodge–Pink lattices) and the corresponding condition on the Hodge polygon.

We are going to show that these stacks are Artin stacks of finite type over $E_\mu$.

Locally on Spec $R$ we may choose an isomorphism $D \cong (R \otimes_{\mathbb{Q}_p} K_0)^d$ by Lemma 2.3(a). Then $\Phi$ and $N$ correspond to matrices $\Phi \in \mathrm{GL}_d(R \otimes_{\mathbb{Q}_p} K_0) = (\mathrm{Res}_{K_0/\mathbb{Q}_p} \mathrm{GL}_{d,K_0})(R)$ and $N \in \mathrm{Mat}_{d\times d}(R \otimes_{\mathbb{Q}_p} K_0) = (\mathrm{Res}_{K_0/\mathbb{Q}_p} \mathrm{Mat}_{d\times d})(R)$. The relation $\Phi\, \varphi^* N = p\, N\, \Phi$ is represented by a closed subscheme

$$P_{K_0,d} \subset (\mathrm{Res}_{K_0/\mathbb{Q}_p} \mathrm{GL}_{d,K_0}) \times_{\operatorname{Spec} \mathbb{Q}_p} (\mathrm{Res}_{K_0/\mathbb{Q}_p} \mathrm{Mat}_{d\times d}).$$

**Theorem 3.2.** (a) *The $\mathbb{Q}_p$-scheme $P_{K_0,d}$ is reduced, Cohen–Macaulay, generically smooth and equidimensional of dimension $f d^2$. In the notation of Remark 2.4 the matrix $(\Phi^f)_0$ has no multiple eigenvalues at the generic points of the irreducible components of $P_{K_0,d}$.*

(b) *The generic points of $P_{K_0,d}$ are in bijection with the partitions $d = k_1 + \cdots + k_m$ for integers $m$ and $1 \le k_1 \le \cdots \le k_m$. To such a partition corresponds the generic point at which the suitably ordered eigenvalues $\lambda_1, \ldots, \lambda_d$ of $(\Phi^f)_0$ satisfy $p^f \lambda_i = \lambda_j$ if and only if $j = i + 1$ and $i \notin \{k_1, k_1 + k_2, \ldots, k_1 + \cdots + k_m\}$. Equivalently to such a partition corresponds the generic point at which the nilpotent endomorphism $N_0$, in the notation of Remark 2.4, has Jordan canonical form with $m$ Jordan blocks of size $k_1, \ldots, k_m$.*

For the proof we will need the following lemma.

**Lemma 3.3.** *Let $r_1, \ldots, r_n$ be integers with $r_1 + \cdots + r_n \ge n$. Then $\sum_{i=1}^n r_i^2 - \sum_{i=1}^{n-1} r_i\, r_{i+1} > 1$, except for the case when $r_1 = \cdots = r_n = 1$.*

*Proof.* We multiply the inequality by 2 and write it as $r_1^2 + \sum_{i=1}^{n-1} (r_i - r_{i+1})^2 + r_n^2 > 2$. There are the following three critical cases:

(a) $\sum_i (r_i - r_{i+1})^2 = 0$,

(b) $\sum_i (r_i - r_{i+1})^2 = 1$,

(c) $\sum_i (r_i - r_{i+1})^2 = 2$.

In case (a) we have $r_1 = \cdots = r_n$. Since $r_1 = \cdots = r_n = 1$ was excluded and $r_1 \le 0$ contradicts $r_1 + \cdots + r_n \ge n$, we have $r_1^2 + r_n^2 > 2$.

In case (b) there is exactly one index $1 \le i < n$ with $r_1 = \cdots = r_i \ne r_{i+1} = \cdots = r_n$ and $|r_i - r_{i+1}| = 1$. If $r_1 \ne 0 \ne r_n$ then $r_1^2 + \sum_{i=1}^{n-1} (r_i - r_{i+1})^2 + r_n^2 > 2$. On the other hand, if $r_1 = \pm 1$ and $r_n = 0$, then $\sum_\nu r_\nu = \pm i < n$. And if $r_1 = 0$ and $r_n = \pm 1$, then $\sum_\nu r_\nu = \pm(n - i) < n$. Both are contradictions.

In case (c) there are exactly two indices $1 \le i < j < n$ with $r_1 = \cdots = r_i$ and $r_{i+1} = \cdots = r_j$ and $r_{j+1} = \cdots = r_n$, as well as $|r_i - r_{i+1}| = 1 = |r_j - r_{j+1}|$. If in addition $r_1 = r_n = 0$ then $\sum_i r_i = \pm(j - i) < n$, which is a contradiction. Therefore $r_1^2 + r_n^2 > 0$ and $r_1^2 + \sum_{i=1}^{n+1} (r_i - r_{i+1})^2 + r_n^2 > 2$ as desired. □

*Proof of Theorem 3.2.* We break the proof into several steps.

1. By [EGA IV$_2$ 1965, Proposition 6.5.3, Corollaires 6.3.5(ii), 6.1.2; EGA IV$_4$ 1967, Proposition 17.7.1] the statement may be checked after the finite étale base change $\operatorname{Spec} K_0 \to \operatorname{Spec} \mathbb{Q}_p$. We will use throughout that after this base change, Remark 2.4 allows to decompose $\Phi = (\Phi_i)_i$ and $N = (N_i)_i$ such that $p\,\Phi_i \circ N_i = N_{i+1} \circ \Phi_i$.

2. We first prove that all irreducible components of $P_{K_0,d}$ have dimension greater or equal to $fd^2$. Sending $(\Phi, N)$ to the entries of the matrices $\Phi_i$, $N_i$ embeds $P_{K_0,d} \times_{\mathbb{Q}_p} K_0$ into affine space $\mathbb{A}^{2fd^2}_{K_0}$ as a locally closed subscheme cut out by the $fd^2$ equations $p\,\Phi_i \circ N_i = N_{i+1} \circ \Phi_i$ for $i = 0, \ldots, f-1$. Therefore the codimension of $P_{K_0,d} \times_{\mathbb{Q}_p} K_0$ in $\mathbb{A}^{2fd^2}_{K_0}$ is less or equal to $fd^2$ by Krull's principal ideal theorem [Eisenbud 1995, Theorem 10.2], and all irreducible components of $P_{K_0,d}$ have dimension greater or equal to $fd^2$ by [Eisenbud 1995, Corollary 13.4].

3. We next prove the assertion on the generic points. Let $y = (\Phi, N)$ be the generic point of an irreducible component $Y$ of $P_{K_0,d}$. After passing to an algebraic closure $L$ of $\kappa(y)$ we may use Remark 2.4 to find a base change matrix $S \in \mathrm{GL}_d(L \otimes_{\mathbb{Q}_p} K_0)$ such that $S^{-1}\Phi\,\varphi(S) = \big((\Phi^f)_0, \mathrm{Id}_d, \ldots, \mathrm{Id}_d\big)$ and $(\Phi^f)_0$ is a block diagonal matrix in Jordan canonical form

$$(\Phi^f)_0 = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_r \end{pmatrix} \quad \text{with } J_i = \begin{pmatrix} \rho_i & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \rho_i \end{pmatrix} \quad \text{and} \quad N_0 = \begin{pmatrix} N_{11} & \cdots & N_{1r} \\ \vdots & & \vdots \\ N_{r1} & \cdots & N_{rr} \end{pmatrix}$$

Note that a priori some of the $\rho_i$ can be equal. Let $s_i$ be the size of the Jordan block $J_i$. Then $N_{ij}$ is an $s_i \times s_j$-matrix. The condition $p^f (\Phi^f)_0 \circ N_0 = N_0 \circ (\Phi^f)_0$ is equivalent to $p^f J_i\, N_{ij} = N_{ij} J_j$ for all $i, j$. It yields $N_{ij} = (0)$ for $p^f \rho_i \neq \rho_j$. By renumbering the $J_i$ we may assume that $N_{ij} \neq (0)$ implies $i < j$. We set $N_{ij} = \big(n^{(ij)}_{\mu,\nu}\big)_{\mu=1\ldots s_i,\, \nu=1\ldots s_j}$. When $p^f \rho_i = \rho_j$ it follows from

$$\begin{pmatrix} p^f n_{2,1} & \cdots & p^f n_{2,s_j} \\ \vdots & & \vdots \\ p^f n_{s_i,1} & \cdots & p^f n_{s_i,s_j} \\ 0 & \cdots & 0 \end{pmatrix} = p^f (J_i - \rho_i) N_{ij} = N_{ij}(J_j - \rho_j) = \begin{pmatrix} 0 & n_{1,1} & \cdots & n_{1,s_j-1} \\ \vdots & \vdots & & \vdots \\ 0 & n_{s_i,1} & \cdots & n_{s_i,s_j-1} \end{pmatrix}$$

that $p^f n^{(ij)}_{\mu,\nu} = n^{(ij)}_{\mu-1,\nu-1}$ for all $\mu, \nu \geq 2$ and $n^{(ij)}_{\mu,\nu} = 0$ whenever $\mu - \nu > \min\{0, s_i - s_j\}$. We set $s := \max\{s_i\}$. The assertion of the theorem says that $s = 1$ and that all $\rho_i$ are pairwise different.

First assume that $s > 1$. We exhibit a morphism $\operatorname{Spec} L[z, z^{-1}] \to P_{K_0,d}$ which sends the point $\{z = 1\}$ to $y$ and the generic point $\operatorname{Spec} L(z)$ to a point at which the maximal size of the Jordan blocks is strictly less than $s$. Since $y$ was a generic point of $P_{K_0,d}$ this is impossible. The morphism $\operatorname{Spec} L[z, z^{-1}] \to P_{K_0,d}$

is given by matrices $\widetilde{S}$, $(\widetilde{\Phi}^f)_0$ and $\widetilde{N}_0$ as follows. We set $\widetilde{S} := S$. For all $i$ with $s_i = s$ we set

$$\begin{pmatrix} \rho_i & 1 & & \\ & \ddots & \ddots & \\ & & \rho_i & 1 \\ & & & z\rho_i \end{pmatrix}$$

and for all $i$ with $s_i < s$ we set $\widetilde{J}_i := J_i$. When $p^f\rho_i \neq \rho_j$ we set $\widetilde{N}_{ij} := (0)$. To define $\widetilde{N}_{ij}$ when $p^f\rho_i = \rho_j$, and hence $i < j$, we distinguish the following cases:

(a) If $s_i, s_j < s$ we set $\widetilde{N}_{ij} = N_{ij}$.

(b) If $s_i = s > s_j$ we set $\widetilde{N}_{ij} = N_{ij}$.

(c) If $s_i < s = s_j$ we set $\widetilde{N}_{ij} = \big(\tilde{n}^{(ij)}_{\mu,\nu}\big)_{\mu,\nu}$ with $\tilde{n}^{(ij)}_{\mu,s_j} := n^{(ij)}_{\mu,s_j}$ for all $\mu$, with $\tilde{n}^{(ij)}_{\mu,\nu} := 0$ whenever $\mu > \nu + s_i - s_j + 1$, and with

$$\tilde{n}^{(ij)}_{\mu,\nu} := n^{(ij)}_{\mu,\nu} + (1-z)p^{(s_j-1-\nu)f} \cdot \rho_j \cdot n^{(ij)}_{\mu-\nu+s_j-1,\,s_j}$$

for $\nu < s_j$ and $\mu \leq \nu + s_i - s_j + 1$.

(d) If $s_i = s_j = s$ we set $\widetilde{N}_{ij} = \big(\tilde{n}^{(ij)}_{\mu,\nu}\big)_{\mu,\nu}$ with $\tilde{n}^{(ij)}_{\mu,s} := n^{(ij)}_{\mu,s}$ for all $\mu$, with $\tilde{n}^{(ij)}_{\mu,\nu} := 0$ whenever $\mu > \nu$, and with $\tilde{n}^{(ij)}_{\mu,\nu} := n^{(ij)}_{\mu,\nu} + (1-z)p^{(s-1-\nu)f} \cdot \rho_j \cdot n^{(ij)}_{\mu-\nu+s-1,\,s}$ for all $\mu \leq \nu < s$.

We have to check that $p^f\widetilde{J}_i\,\widetilde{N}_{ij} = \widetilde{N}_{ij}\widetilde{J}_j$ for all $i$, $j$ with $p^f\rho_i = \rho_j$. In case (a) this is obvious and in case (b) it follows from the fact that the bottom row of $N_{ij}$ is zero. For case (c) we compute

$$p^f(\widetilde{J}_i - \rho_i)\widetilde{N}_{ij} = \begin{pmatrix} p^f\tilde{n}_{2,1} & \cdots & p^f\tilde{n}_{2,s_j} \\ \vdots & & \vdots \\ p^f\tilde{n}_{s_i,1} & \cdots & p^f\tilde{n}_{s_i,s_j} \\ 0 & \cdots & 0 \end{pmatrix} = \begin{pmatrix} 0 & \tilde{n}_{1,1} & \cdots & \tilde{n}_{1,s_j-2} & \tilde{n}_{1,s_j-1} + (z-1)\rho_j\tilde{n}_{1,s_j} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & \tilde{n}_{s_i,1} & \cdots & \tilde{n}_{s_i,s_j-2} & \tilde{n}_{s_i,s_j-1} + (z-1)\rho_j\tilde{n}_{s_i,s_j} \end{pmatrix}$$

$$= \widetilde{N}_{ij}(\widetilde{J}_j - \rho_j).$$

Finally for case (d) we compute

$$p^f(\widetilde{J}_i - \rho_i)\widetilde{N}_{ij} = \begin{pmatrix} p^f\tilde{n}_{2,1} & \cdots & p^f\tilde{n}_{2,s-1} & p^f\tilde{n}_{2,s} \\ \vdots & & \vdots & \vdots \\ p^f\tilde{n}_{s,1} & \cdots & p^f\tilde{n}_{s,s-1} & p^f\tilde{n}_{s,s} \\ 0 & \cdots & 0 & (z-1)p^f\rho_i\tilde{n}_{s,s} \end{pmatrix}$$

$$= \begin{pmatrix} 0 & \tilde{n}_{1,1} & \cdots & \tilde{n}_{1,s-2} & \tilde{n}_{1,s-1} + (z-1)\rho_j\tilde{n}_{1,s} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & \tilde{n}_{s,1} & \cdots & \tilde{n}_{s,s-2} & \tilde{n}_{s,s-1} + (z-1)\rho_j\tilde{n}_{s,s} \end{pmatrix}$$

$$= \widetilde{N}_{ij}(\widetilde{J}_j - \rho_j).$$

Altogether this defines the desired morphism $\operatorname{Spec} L[z, z^{-1}] \to P_{K_0,d}$.

So we have shown that $s = 1$ at the generic point $y$ and that $(\Phi^f)_0$ is a diagonal matrix. We still have to show that all diagonal entries are pairwise different. For this purpose we rewrite $(\Phi^f)_0$ and $N_0$ as

$$(\Phi^f)_0 = \begin{pmatrix} \lambda_1 \operatorname{Id}_{r_1} & & \\ & \ddots & \\ & & \lambda_n \operatorname{Id}_{r_n} \end{pmatrix} \qquad \text{and} \qquad N_0 = \begin{pmatrix} M_{11} & \cdots & M_{1n} \\ \vdots & & \vdots \\ M_{n1} & \cdots & M_{nn} \end{pmatrix}$$

We denote the multiplicity of the eigenvalue $\lambda_i$ by $r_i \geq 1$. Then $M_{ij}$ is an $r_i \times r_j$-matrix. By renumbering the $\lambda_i$ we may assume that there are indices $0 = l_0 < l_1 < \cdots < l_m = d$ such that $p^f \lambda_i = \lambda_j$ if and only if $j = i + 1$ and $i \notin \{l_1, \ldots, l_m\}$.

We compute $\dim Y = \operatorname{trdeg}_{\mathbb{Q}_p} \kappa(y) = \operatorname{trdeg}_{\mathbb{Q}_p} L$ as follows. The eigenvalues $\lambda_{l_1}, \ldots, \lambda_{l_m}$ contribute at most the summand $m$ to $\operatorname{trdeg}_{\mathbb{Q}_p} L$.

The matrix $S \in \mathrm{GL}_d(L \otimes_{\mathbb{Q}_p} K_0)$ is determined only up to multiplication on the right with an element of the $\varphi$-centralizer

$$\mathcal{C}(L) := \left\{ S \in \mathrm{GL}_d(L \otimes_{\mathbb{Q}_p} K_0) : S\big((\Phi^f)_0, \operatorname{Id}_d, \ldots, \operatorname{Id}_d\big) = \big((\Phi^f)_0, \operatorname{Id}_d, \ldots, \operatorname{Id}_d\big)\varphi(S) \right\}$$

of $\big((\Phi^f)_0, \operatorname{Id}_d, \ldots, \operatorname{Id}_d\big)$. Writing $S = (S_0, \ldots, S_{f-1})$ this condition implies that $S_i \overset{!}{=} (\varphi(S))_i := S_{i-1}$ for $i = 1, \ldots, f-1$ and $S_0(\Phi^f)_0 \overset{!}{=} (\Phi^f)_0(\varphi(S))_0 := (\Phi^f)_0 S_{f-1} = (\Phi^f)_0 S_0$. Therefore $\mathcal{C}$ has dimension $\sum_i r_i^2$ and the entries of $S \in (\operatorname{Res}_{K_0/\mathbb{Q}_p} \mathrm{GL}_{d,K_0})/\mathcal{C}$ contribute at most the summand $fd^2 - \sum_i r_i^2$ to $\operatorname{trdeg}_{\mathbb{Q}_p} L$.

The condition $p^f (\Phi^f)_0 \circ N_0 = N_0 \circ (\Phi^f)_0$ is equivalent to $p^f \lambda_i M_{ij} = \lambda_j M_{ij}$ for all $i, j$. This implies that there is no condition on $M_{ij}$ when $j = i + 1$ and $i \notin \{l_1, \ldots, l_m\}$, and that all other $M_{ij}$ are zero. So the entries of the $M_{ij}$ contribute at most the summand $\sum_{i \notin \{l_1, \ldots, l_m\}} r_i\, r_{i+1}$ to $\operatorname{trdeg}_{\mathbb{Q}_p} L$.

Adding all summands and comparing with our estimate in part 2 above, we obtain

$$\begin{aligned} fd^2 \leq \dim Y = \operatorname{trdeg}_{\mathbb{Q}_p} \kappa(y) &\leq m + fd^2 - \sum_{i=1}^{n} r_i^2 + \sum_{i \notin \{l_1, \ldots, l_m\}} r_i\, r_{i+1} \\ &= fd^2 + \sum_{\nu=0}^{m-1} \Bigg( 1 - \sum_{i=1+l_\nu}^{l_{\nu+1}} r_i^2 + \sum_{i=1+l_\nu}^{l_{\nu+1}-1} r_i\, r_{i+1} \Bigg). \end{aligned}$$

By Lemma 3.3 the parentheses are zero when all $r_i = 1$, and negative otherwise. So we have proved that $r_1 = \cdots = r_n = 1$. In other words, all diagonal entries of $(\Phi^f)_0$ are pairwise different. Let $k_\nu := l_\nu - l_{\nu-1}$ for $\nu = 1, \ldots, m$. Then the generic point $y$ corresponds to the partition $d = k_1 + \cdots + k_m$ under the description of the generic points in the theorem. As we have noticed above the $1 \times 1$ matrices $M_{ij}$ vanish at $y$ unless $j = i + 1$ and $i \notin \{l_1, \ldots, l_m\}$ and in the latter case we must have $M_{ij}(y) \neq 0$. This implies the claim on the Jordan type of $N_0$ at the generic points of the irreducible components.

Moreover, it follows that $\dim Y = fd^2$ for all irreducible components $Y$ of $P_{K_0,d}$. By [Eisenbud 1995, Proposition 18.13] this also implies that $P_{K_0,d} \times_{\mathbb{Q}_p} K_0$ is Cohen–Macaulay.

4. It remains to show that $P_{K_0,d}$ is generically smooth over $\mathbb{Q}_p$. From this it follows that it is reduced, because it is Cohen–Macaulay. Let again $y$ be the generic point of an irreducible component of $P_{K_0,d}\times_{\mathbb{Q}_p} K_0$ and let $L$ be an algebraic closure of $\kappa(y)$. As above, Remark 2.4 allows us to change the basis over $L$ and assume that $\Phi=((\Phi_0^f),\ \mathrm{id},\ldots,\ \mathrm{id})$ and $N=(p^iN_0)_i$ with $(\Phi^f)_0=\mathrm{diag}(\lambda_1,\ldots,\lambda_d)$ and $\lambda_i\neq\lambda_j$ for all $i\neq j$. We write $F^{(0)}:=(\Phi^f)_0$ and $N_0^{(0)}:=N_0=(n_{ij})_{ij}$. The condition $N_0^{(0)}F^{(0)}=p^fF^{(0)}N_0^{(0)}$ implies that $n_{ij}=0$ if $p^f\lambda_i\neq\lambda_j$. And conversely $n_{ij}\neq 0$ if $p^f\lambda_i=\lambda_j$ by our explicit description of $N_0$ at $y$ above.

We claim that for every $n\geq 1$, any deformation $(F^{(n-1)},N_0^{(n-1)})\in P_{K_0,d}(L[\varepsilon]/\varepsilon^n)$ of $(F^{(0)},N_0^{(0)})$ can be lifted further to $(F^{(n)},N_0^{(n)})\in P_{K_0,d}(L[\varepsilon]/\varepsilon^{n+1})$. This implies that $P_{K_0,d}$ is smooth at $y$, as it follows that any tangent vector $\mathcal{O}_{P_{K_0,d},y}\to L[\varepsilon]/\varepsilon^2$ comes from a map $\mathcal{O}_{P_{K_0,d},y}\to L[\![\varepsilon]\!]$ and hence the image of

$$\mathrm{Spec}\Big(\bigoplus\nolimits_{i\geq 0}(\mathfrak{m}^i_{P_{K_0,d},y}/\mathfrak{m}^{i+1}_{P_{K_0,d},y})\Big)\to\mathrm{Spec}\left(\mathrm{Sym}^\bullet(\mathfrak{m}_{P_{K_0,d},y}/\mathfrak{m}^2_{P_{K_0,d},y})\right)$$

contains any tangent vector. This means that the tangent cone at $y$ equals the tangent space, and hence by [Mumford 1999, III, §4 Definition 2 and Corollary 1] $P_{K_0,d}$ is smooth at $y$. Let us take any deformation $(\widetilde{F},\widetilde{N}_0)\in\mathrm{GL}_d(L[\varepsilon]/\varepsilon^{n+1})\times\mathrm{Mat}_{d\times d}(L[\varepsilon]/\varepsilon^{n+1})$ of $(F^{(n-1)},N_0^{(n-1)})$. Then we have $\widetilde{N}_0\widetilde{F}-p^f\widetilde{F}\widetilde{N}_0\in\varepsilon^n\mathrm{Mat}_{d\times d}(L)$. Changing $(\widetilde{F},\widetilde{N}_0)$ to $(F^{(n)},N_0^{(n)})=(\widetilde{F}+\varepsilon^nF',\widetilde{N}_0+\varepsilon^nN_0')$ with $F',N_0'\in\mathrm{Mat}_{d\times d}(L)$ we find

$$N_0^{(n)}F^{(n)}-p^fF^{(n)}N_0^{(n)}=(\widetilde{N}_0\widetilde{F}-p^f\widetilde{F}\widetilde{N}_0)+\varepsilon^n(N_0^{(0)}F'+N_0'F^{(0)}-p^f(F'N_0^{(0)}+F^{(0)}N_0'))$$

and hence it suffices to show that the map $h:\mathrm{Mat}_{d\times d}(L)\times\mathrm{Mat}_{d\times d}(L)\to\mathrm{Mat}_{d\times d}(L)$ given by

$$h:(F',N_0')\mapsto(N_0'F^{(0)}-p^fF^{(0)}N_0')+(N_0^{(0)}F'-p^fF'N_0^{(0)})$$

is surjective. For this purpose let $N_0'=(b_{ij})_{ij}$ and $F'=\mathrm{diag}(a_1,\ldots,a_d)$. Then we find that

$$h_{ij}(F',N_0'):=h(F',N_0')_{ij}=(\lambda_j-p^f\lambda_i)b_{ij}+a_j\,n_{ij}-p^fa_i\,n_{ij}.$$

Whenever $\lambda_j-p^f\lambda_i\in L^\times$ and hence $n_{ij}=0$ we obtain the surjectivity of $h_{ij}$. By permuting the indices we may assume that $n_{ij}\neq 0$ implies $j=i+1$. Treating every Jordan block of $N_0^{(0)}$ separately we may further assume that $n_{i,i+1}\neq 0$ for all $i$. It then follows that we have

$$h_{i,i+1}(F',N_0')=(a_{i+1}-p^fa_i)n_{i,i+1}$$

which suffices to see that the map $h$ is surjective. □

**Remark 3.4.** The scheme $P_{K_0,d}$ is in general not normal. For example if $K=\mathbb{Q}_p$ and $d=2$ then $P_{\mathbb{Q}_p,2}$ has two generic points. This was already proven in [Kisin 2009, Lemma A.3]. In one of the generic points $\Phi$ has eigenvalues $\lambda$, $p\lambda$ and $N\neq 0$. In the other $\Phi$ has eigenvalues $\lambda_1,\lambda_2$ with $\lambda_j\neq p\lambda_i$ for all $i,j$ and $N=0$. Both irreducible components meet in the codimension one point where $\lambda_2=p\lambda_1$ and $N=0$.

Let $\Delta$ denote the set of simple roots (defined over $\overline{\mathbb{Q}}_p$) of $\widetilde{G}:=\mathrm{Res}_{K/\mathbb{Q}_p}\mathrm{GL}_{d,K}$ with respect to the Borel subgroup $\widetilde{B}$ and denote by $\Delta_\mu\subset\Delta$ the set of all simple roots $\alpha$ such that $\langle\alpha,\mu\rangle=0$. Here $\langle-,-\rangle$ is the canonical pairing between characters and cocharacters. We write $P_\mu$ for the parabolic subgroup

of $\widetilde{G}$ containing $\widetilde{B}$ and corresponding to $\Delta_\mu \subset \Delta$. This parabolic subgroup is defined over $E_\mu$, and the quotient by this parabolic is a projective $E_\mu$-variety

$$\mathrm{Flag}_{K,d,\mu} = \widetilde{G}_{E_\mu}/P_\mu \tag{3-1}$$

representing the functor

$$R \mapsto \big\{\text{filtrations } \mathcal{F}^\bullet \text{ of } R\otimes_{\mathbb{Q}_p} K^{\oplus d} \text{ with constant Hodge polygon equal to } \mu\big\}$$

Thus $\mathscr{D}_{\varphi,N,\mu}$ and $\mathscr{D}_{\varphi,\mu}$ are isomorphic to the stack quotients

$$\begin{aligned}\mathscr{D}_{\varphi,N,\mu} &\cong (P_{K_0,d} \times_{\mathrm{Spec}\,\mathbb{Q}_p} \mathrm{Flag}_{K,d,\mu}) \,\big/\, (\mathrm{Res}_{K_0/\mathbb{Q}_p}\mathrm{GL}_{d,K_0})_{E_\mu},\\ \mathscr{D}_{\varphi,\mu} &\cong (\mathrm{Res}_{K_0/\mathbb{Q}_p}\mathrm{GL}_{d,K_0} \times_{\mathrm{Spec}\,\mathbb{Q}_p} \mathrm{Flag}_{K,d,\mu}) \,\big/\, (\mathrm{Res}_{K_0/\mathbb{Q}_p}\mathrm{GL}_{d,K_0})_{E_\mu},\end{aligned} \tag{3-2}$$

where $g \in (\mathrm{Res}_{K_0/\mathbb{Q}_p}\mathrm{GL}_{d,K_0})_{E_\mu}$ acts on $(\Phi, N, \mathcal{F}^\bullet) \in P_{K_0,d} \times_{\mathrm{Spec}\,\mathbb{Q}_p} \mathrm{Flag}_{K,d,\mu}$ by

$$(\Phi, N, \mathcal{F}^\bullet) \mapsto \big(g^{-1}\Phi\,\varphi(g),\, g^{-1}Ng,\, g^{-1}\mathcal{F}^\bullet\big).$$

We next describe the moduli space for the Hodge–Pink lattice $\mathfrak{q}$. Fix integers $m = \max\{\mu_{\psi,1} \mid \psi : K \to \widetilde{K}\}$ and $n = \max\{-\mu_{\psi,d} \mid \psi : K \to \widetilde{K}\}$. Then by Cramer's rule we have $E(u)^n\mathfrak{p} \subset \mathfrak{q} \subset E(u)^{-m}\mathfrak{p}$. So $\mathfrak{q}$ is determined by the epimorphism

$$pr : R\otimes_{\mathbb{Q}_p} (K[t]/t^{m+n})^{\oplus d} \xrightarrow{\sim} E(u)^{-m}\mathfrak{p}/E(u)^n\mathfrak{p} \twoheadrightarrow E(u)^{-m}\mathfrak{p}/\mathfrak{q} \tag{3-3}$$

which is induced by choosing an isomorphism $D \cong (R\otimes_{\mathbb{Q}_p} K_0)^d$ locally on $R$. The quotient $E(u)^{-m}\mathfrak{p}/\mathfrak{q}$ is a finite locally free $R$-module and of finite presentation over $R\otimes_{\mathbb{Q}_p} K[t]/t^{m+n}$ by Lemma 2.7. Therefore it is an $R$-valued point of Grothendieck's Quot-scheme $\mathrm{Quot}_{\mathcal{O}^d \mid K[t]/t^{m+n} \mid \mathbb{Q}_p}$; see [Grothendieck 1962, n°221, Theorem 3.1] or [Altman and Kleiman 1980, Theorem 2.6]. This Quot-scheme is projective over $\mathbb{Q}_p$. The boundedness by $\mu$ is represented by a closed subscheme $Q_{K,d,\preceq\mu}$ of $\mathrm{Quot}_{\mathcal{O}^d \mid K[t]/t^{m+n} \mid \mathbb{Q}_p} \times_{\mathrm{Spec}\,\mathbb{Q}_p} \mathrm{Spec}\,E_\mu$ according to Proposition 2.13(a). Thus $\mathscr{H}_{\varphi,N,\preceq\mu}$ and $\mathscr{H}_{\varphi,\preceq\mu}$ are isomorphic to the stack quotients

$$\begin{aligned}\mathscr{H}_{\varphi,N,\preceq\mu} &\cong (P_{K_0,d} \times_{\mathrm{Spec}\,\mathbb{Q}_p} Q_{K,d,\preceq\mu}) \,\big/\, (\mathrm{Res}_{K_0/\mathbb{Q}_p}\mathrm{GL}_{d,K_0})_{E_\mu},\\ \mathscr{H}_{\varphi,\preceq\mu} &\cong (\mathrm{Res}_{K_0/\mathbb{Q}_p}\mathrm{GL}_{d,K_0} \times_{\mathrm{Spec}\,\mathbb{Q}_p} Q_{K,d,\preceq\mu}) \,\big/\, (\mathrm{Res}_{K_0/\mathbb{Q}_p}\mathrm{GL}_{d,K_0})_{E_\mu},\end{aligned}$$

where $g \in (\mathrm{Res}_{K_0/\mathbb{Q}_p}\mathrm{GL}_{d,K_0})_{E_\mu}$ acts on $(\Phi, N, pr) \in P_{K_0,d} \times_{\mathrm{Spec}\,\mathbb{Q}_p} Q_{K,d,\preceq\mu}$ with $pr$ from (3-3) by

$$(\Phi, N, pr) \mapsto \big(g^{-1}\Phi\,\varphi(g),\ g^{-1}Ng,\ pr\circ(g\otimes_{\mathbb{Q}_p} \mathrm{id}_{\mathbb{Q}_p[t]/t^{m+n}})\big).$$

Let $Q_{K,d,\mu}$ be the complement in $Q_{K,d,\preceq\mu}$ of the image of

$$\bigcup_{\mu'\prec\mu} Q_{K,d,\preceq\mu'} \times_{\mathrm{Spec}\,E_{\mu'}} \mathrm{Spec}\,\widetilde{K}$$

under the finite étale projection $Q_{K,d,\preceq\mu} \times_{\mathrm{Spec}\,E_\mu} \mathrm{Spec}\,\widetilde{K} \to Q_{K,d,\preceq\mu}$. Here the union is taken over all dominant cocharacters $\mu' : \mathbb{G}_{m,\overline{\mathbb{Q}}_p} \to \widetilde{T}_{\overline{\mathbb{Q}}_p}$ which are strictly less than $\mu$ in the Bruhat order; see Proposition 2.13(b). Since there are only finitely many such $\mu'$ the scheme $Q_{K,d,\mu}$ is an open subscheme

of $Q_{K,d,\preceq\mu}$ and quasiprojective over $E_\mu$. By Proposition 2.13(a) the stacks $\mathscr{H}_{\varphi,N,\mu} \subset \mathscr{H}_{\varphi,N,\preceq\mu}$ and $\mathscr{H}_{\varphi,\mu} \subset \mathscr{H}_{\varphi,\preceq\mu}$ are therefore open substacks and isomorphic to the stack quotients

$$\mathscr{H}_{\varphi,N,\mu} \cong (P_{K_0,d} \times_{\operatorname{Spec}\mathbb{Q}_p} Q_{K,d,\mu}) \,/\, (\operatorname{Res}_{K_0/\mathbb{Q}_p} \mathrm{GL}_{d,K_0})_{E_\mu},$$
$$\mathscr{H}_{\varphi,\mu} \cong (\operatorname{Res}_{K_0/\mathbb{Q}_p} \mathrm{GL}_{d,K_0} \times_{\operatorname{Spec}\mathbb{Q}_p} Q_{K,d,\mu}) \,/\, (\operatorname{Res}_{K_0/\mathbb{Q}_p} \mathrm{GL}_{d,K_0})_{E_\mu}.$$

There is another description of $Q_{K,d,\preceq\mu}$ in terms of the affine Grassmannian. Consider the infinite-dimensional affine group schemes $L^+\mathrm{GL}_d$ and $L^+\widetilde{G}$ over $\mathbb{Q}_p$, and the sheaves $L\mathrm{GL}_d$ and $L\widetilde{G}$ for the fpqc-topology on $\mathbb{Q}_p$ whose sections over a $\mathbb{Q}_p$-algebra $R$ are given by

$$\begin{aligned}
L^+\mathrm{GL}_d(R) &= \mathrm{GL}_d(R[\![t]\!]),\\
L^+\widetilde{G}(R) &= \widetilde{G}(R[\![t]\!]) = \mathrm{GL}_d(R \otimes_{\mathbb{Q}_p} K[\![t]\!]) = \mathrm{GL}_d(\mathbb{B}_R^+),\\
L\mathrm{GL}_d(R) &= \mathrm{GL}_d\big(R[\![t]\!]\big[\tfrac{1}{t}\big]\big),\\
L\widetilde{G}(R) &= \widetilde{G}\big(R[\![t]\!]\big[\tfrac{1}{t}\big]\big) = \mathrm{GL}_d\big(R \otimes_{\mathbb{Q}_p} K[\![t]\!]\big[\tfrac{1}{t}\big]\big) = \mathrm{GL}_d(\mathbb{B}_R).
\end{aligned}$$

$L^+\mathrm{GL}_d$ and $L^+\widetilde{G}$ are called the *group of positive loops*, and $L\mathrm{GL}_d$ and $L\widetilde{G}$ are called the *loop group* of $\mathrm{GL}_d$ (resp. $\widetilde{G}$). The *affine Grassmannian* of $\mathrm{GL}_d$ (resp. $\widetilde{G}$) is the quotient sheaf for the fppf-topology on $\mathbb{Q}_p$

$$\mathrm{Gr}_{\mathrm{GL}_d} := L\mathrm{GL}_d\,/L^+\mathrm{GL}_d \qquad (\text{resp. } \mathrm{Gr}_{\widetilde{G}} := L\widetilde{G}/L^+\widetilde{G}).$$

They are ind-schemes over $\mathbb{Q}_p$ which are ind-projective; see [Beilinson and Drinfeld 2005, §4.5; Beauville and Laszlo 1994; Laszlo and Sorger 1997; Hartl and Viehmann 2011].

We set $\mathrm{Gr}_{\mathrm{GL}_d,\widetilde{K}} := \mathrm{Gr}_{\mathrm{GL}_d} \times_{\operatorname{Spec}\mathbb{Q}_p} \operatorname{Spec}\widetilde{K}$. Then there are morphisms

$$\begin{aligned}
Q_{K,d,\preceq\mu} &\to \mathrm{Gr}_{\widetilde{G}} \times_{\operatorname{Spec}\mathbb{Q}_p} \operatorname{Spec} E_\mu =: \mathrm{Gr}_{\widetilde{G},E_\mu},\\
Q_{K,d,\preceq\mu} \times_{\operatorname{Spec} E_\mu} \operatorname{Spec}\widetilde{K} &\to \prod_{\psi:K\to\widetilde{K}} \mathrm{Gr}_{\mathrm{GL}_d,\widetilde{K}},
\end{aligned} \tag{3-4}$$

which are defined as follows. Let $\mathfrak{q} \subset (\mathbb{B}_{Q_{K,d,\preceq\mu}})^{\oplus d}$ be the universal Hodge–Pink lattice over $Q_{K,d,\preceq\mu}$. Then by Lemma 2.7 there is an étale covering $f : \operatorname{Spec} R \to Q_{K,d,\preceq\mu}$ such that $f^*\mathfrak{q}$ is free over $\mathbb{B}_R^+$. With respect to a basis of $f^*\mathfrak{q}$ the equality $\mathbb{B}_R \cdot f^*\mathfrak{q} = D \otimes_{R\otimes K_0} \mathbb{B}_R$ corresponds to a matrix $A \in \mathrm{GL}_d(\mathbb{B}_R) = L\widetilde{G}(R)$. The image of $A$ in $\mathrm{Gr}_{\widetilde{G}}(R)$ is independent of the basis and by étale descend defines the first factor of the map $Q_{K,d,\preceq\mu} \to \mathrm{Gr}_{\widetilde{G}} \times_{\operatorname{Spec}\mathbb{Q}_p} \operatorname{Spec} E_\mu$. The base change of this map along the finite étale morphism $\operatorname{Spec}\widetilde{K} \to \operatorname{Spec} E_\mu$ defines the second map in (3-4), using the splitting $\widetilde{G} \times_{\mathbb{Q}_p} \widetilde{K} = \prod_\psi \mathrm{GL}_{d,\widetilde{K}}$ which induces similar splittings for $L^+\widetilde{G}$, $L\widetilde{G}$, and $\mathrm{Gr}_{\widetilde{G}}$.

The boundedness by $\mu$ is represented by closed ind-subschemes

$$\mathrm{Gr}^{\preceq\mu}_{\widetilde{G},E_\mu} \qquad \text{and} \qquad \mathrm{Gr}^{\preceq\mu}_{\widetilde{G},E_\mu} \times_{\operatorname{Spec} E_\mu} \operatorname{Spec}\widetilde{K} = \prod_\psi \mathrm{Gr}^{\preceq\mu_\psi}_{\mathrm{GL}_d,\widetilde{K}}$$

of $\mathrm{Gr}_{\widetilde{G},E_\mu}$ and $\prod_\psi \mathrm{Gr}_{\mathrm{GL}_d,\widetilde{K}}$, respectively, through which the maps (3-4) factor. Conversely the universal matrix $A$ over $L\widetilde{G}$ defines a $\mathbb{B}^+_{L\widetilde{G}}$-lattice $\mathfrak{q} = A \cdot (\mathbb{B}^+_{L\widetilde{G}})^d$. Its restriction to $\mathrm{Gr}^{\preceq\mu}_{\widetilde{G},E_\mu}$ has Hodge polygon

bounded by $\mu$ and corresponds to the inverses of the maps (3-4). This yields canonical isomorphisms $Q_{K,d,\preceq\mu} \cong \mathrm{Gr}^{\preceq\mu}_{\widetilde{G},E_\mu}$ and $Q_{K,d,\preceq\mu} \times_{\mathrm{Spec}\, E_\mu} \mathrm{Spec}\, \widetilde{K} \cong \prod_\psi \mathrm{Gr}^{\preceq\mu_\psi}_{\mathrm{GL}_d,\widetilde{K}}$. These isomorphisms restrict to isomorphisms of open subschemes $Q_{K,d,\mu} \cong \mathrm{Gr}^{\mu}_{\widetilde{G},E_\mu}$ and $Q_{K,d,\mu} \times_{\mathrm{Spec}\, E_\mu} \mathrm{Spec}\, \widetilde{K} \cong \prod_\psi \mathrm{Gr}^{\mu_\psi}_{\mathrm{GL}_d,\widetilde{K}}$.

In view of [Hartl and Viehmann 2011, §4], especially Lemma 4.3, the boundedness by $\mu$ on $\prod_\psi \mathrm{Gr}^{\preceq\mu_\psi}_{\mathrm{GL}_d,\widetilde{K}}$ can be phrased in terms of Weyl module representations of $\mathrm{GL}_{d,\widetilde{K}}$. In this formulation it was proved by Varshavsky [2004, Proposition A.9] that $\mathrm{Gr}^{\mu_\psi}_{\mathrm{GL}_d,\widetilde{K}}$ is reduced. Therefore this locally closed subscheme is determined by its underlying set of points. Reasoning with the elementary divisor theorem as in Construction 2.10 shows that $\mathrm{Gr}^{\mu_\psi}_{\mathrm{GL}_d,\widetilde{K}}$ is equal to the locally closed Schubert cell

$$L^+\mathrm{GL}_{d,\widetilde{K}} \cdot \mu_\psi(t)^{-1} \cdot L^+\mathrm{GL}_{d,\widetilde{K}} \,\big/\, L^+\mathrm{GL}_{d,\widetilde{K}}$$

and is a homogeneous space under $L^+\mathrm{GL}_{d,\widetilde{K}}$. This description descends to $Q_{K,d,\mu}$ and shows that the latter is reduced and isomorphic to the locally closed Schubert cell $L^+\widetilde{G}_{E_\mu} \cdot \mu(t)^{-1} \cdot L^+\widetilde{G}_{E_\mu} \,/\, L^+\widetilde{G}_{E_\mu}$ which is a homogeneous space under $L^+\widetilde{G}_{E_\mu} := L^+\widetilde{G} \times_{\mathrm{Spec}\, \mathbb{Q}_p} \mathrm{Spec}\, E_\mu$.

These homogeneous spaces can be described more explicitly. Set

$$\begin{aligned} S_{\mathrm{GL}_d,\mu_\psi} &:= L^+\mathrm{GL}_{d,\widetilde{K}} \cap \mu_\psi(t)^{-1} \cdot L^+\mathrm{GL}_{d,\widetilde{K}} \cdot \mu_\psi(t) \subset L^+\mathrm{GL}_{d,\widetilde{K}}, \\ S_{\widetilde{G},\mu} &:= L^+\widetilde{G}_{E_\mu} \cap \mu(t)^{-1} \cdot L^+\widetilde{G}_{E_\mu} \cdot \mu(t) \subset L^+\widetilde{G}_{E_\mu}. \end{aligned}$$

These are closed subgroup schemes and the homogeneous spaces are isomorphic to the quotients

$$\begin{aligned} L^+\mathrm{GL}_{d,\widetilde{K}} / S_{\mathrm{GL}_d,\mu_\psi} &\xrightarrow{\sim} L^+\mathrm{GL}_{d,\widetilde{K}} \cdot \mu_\psi(t)^{-1} \cdot L^+\mathrm{GL}_{d,\widetilde{K}} \,\big/\, L^+\mathrm{GL}_{d,\widetilde{K}}, \\ L^+\widetilde{G}_{E_\mu} / S_{\widetilde{G},\mu} &\xrightarrow{\sim} L^+\widetilde{G}_{E_\mu} \cdot \mu(t)^{-1} \cdot L^+\widetilde{G}_{E_\mu} \,/\, L^+\widetilde{G}_{E_\mu} \cong Q_{K,d,\mu}. \end{aligned}$$

Consider the closed normal subgroup $L^{++}\widetilde{G}_{E_\mu}(R) := \{A \in L^+\widetilde{G}_{E_\mu}(R) : A \equiv 1 \bmod t\}$. Then the parabolic subgroup $P_\mu$ from (3-1) equals

$$P_\mu = S_{\widetilde{G},\mu} \cdot L^{++}\widetilde{G}_{E_\mu} \,/\, L^{++}\widetilde{G}_{E_\mu} \subset L^+\widetilde{G}_{E_\mu} \,/\, L^{++}\widetilde{G}_{E_\mu} = \widetilde{G}_{E_\mu}$$

and this yields a morphism

$$Q_{K,d,\mu} = L^+\widetilde{G}_{E_\mu}/S_{\widetilde{G},\mu} \twoheadrightarrow L^+\widetilde{G}_{E_\mu}/S_{\widetilde{G},\mu} \cdot L^{++}\widetilde{G}_{E_\mu} = \widetilde{G}_{E_\mu}/P_\mu = \mathrm{Flag}_{K,d,\mu}, \tag{3-5}$$

with fibers isomorphic to $S_{\widetilde{G},\mu} \cdot L^{++}\widetilde{G}_{E_\mu}/S_{\widetilde{G},\mu}$. The latter is an affine space because we may consider the base change from $E_\mu$ to $\widetilde{K}$ and the decomposition

$$\big(S_{\widetilde{G},\mu} \cdot L^{++}\widetilde{G}_{E_\mu}/S_{\widetilde{G},\mu}\big) \times_{\mathrm{Spec}\, E_\mu} \mathrm{Spec}\, \widetilde{K} = \prod_\psi \big(S_{\mathrm{GL}_d,\mu_\psi} \cdot L^{++}\,\mathrm{GL}_{d,\widetilde{K}} / S_{\mathrm{GL}_d,\mu_\psi}\big).$$

Each component is an affine space whose $R$-valued points are in bijection with the matrices

$$\begin{pmatrix} 1 & & & \\ a_{21} & 1 & & \\ \vdots & \ddots & \ddots & \\ a_{d1} & \cdots & a_{d,d-1} & 1 \end{pmatrix}$$

where $a_{ij} \in \bigoplus_{k=1}^{\mu_{\psi,j}-\mu_{\psi,i}-1} t^k R$. The Galois group $\mathrm{Gal}(\widetilde{K}/E_\mu)$ canonically identifies the components with the same values for $\mu_\psi$. Therefore $S_{\widetilde{G},\mu} \cdot L^{++}\widetilde{G}_{E_\mu}/S_{\widetilde{G},\mu}$ is an affine space.

We show that $Q_{K,d,\mu}$ is a geometric vector bundle over $\mathrm{Flag}_{K,d,\mu}$ by exhibiting its zero section. The projection $L^+\widetilde{G}_{E_\mu} \twoheadrightarrow \widetilde{G}_{E_\mu}$ has a section given on $R$-valued points by the map $\widetilde{G}_{E_\mu}(R) \to L^+\widetilde{G}_{E_\mu}(R) = \widetilde{G}_{E_\mu}(R[\![t]\!])$ induced from the natural inclusion $R \hookrightarrow R[\![t]\!]$. Since $L^+P_\mu \subset S_{\widetilde{G},\mu}$ by definition of $P_\mu$, this section induces a section

$$\mathrm{Flag}_{K,d,\mu} \;\to\; L^+\widetilde{G}_{E_\mu}/L^+P_\mu \;\to\; L^+\widetilde{G}_{E_\mu}/S_{\widetilde{G},\mu} = Q_{K,d,\mu}.$$

This is the zero section of the geometric vector bundle $Q_{K,d,\mu}$ over $\mathrm{Flag}_{K,d,\mu}$. Using lattices the section coincides (on $L$-valued points for a field $L$) with the map $\mathcal{F}^\bullet \mapsto \mathfrak{q}(\mathcal{F}^\bullet)$ defined in Remark 2.8 (3) and the projection $Q_{K,d,\mu} \to \mathrm{Flag}_{K,d,\mu}$ coincides with the map $\mathfrak{q} \mapsto \mathcal{F}^\bullet_{\mathfrak{q}}$ from Remark 2.8 (1). Let us summarize.

**Proposition 3.5.** (a) *$Q_{K,d,\preceq\mu}$ is projective over $E_\mu$ of dimension $\sum_{\psi,j}(d+1-2j)\mu_{\psi,j}$ and contains $Q_{K,d,\mu}$ as a dense open subscheme. Both schemes are irreducible.*

(b) *$Q_{K,d,\mu}$ is smooth over $E_\mu$ and isomorphic to the homogeneous space $L^+\widetilde{G}_{E_\mu}/S_{\widetilde{G},\mu}$ which is a geometric vector bundle over* $\mathrm{Flag}_{K,d,\mu}$.

*Proof.* Everything was proved above, except the formula for the dimension and the density of $Q_{K,d,\mu}$ which follow from [Beilinson and Drinfeld 2005, 4.5.8, 4.5.12]. The irreducibility of $Q_{K,d,\preceq\mu}$ is a consequence of the density statement. □

**Theorem 3.6.** (a) *The moduli stacks $\mathscr{D}_{\varphi,N,\mu}$, $\mathscr{D}_{\varphi,\mu}$, $\mathscr{H}_{\varphi,N,\preceq\mu}$, $\mathscr{H}_{\varphi,\preceq\mu}$, $\mathscr{H}_{\varphi,N,\mu}$ and $\mathscr{H}_{\varphi,\mu}$ are noetherian Artin stacks of finite type over $E_\mu$.*

(b) *The stack $\mathscr{H}_{\varphi,N,\mu}$ is a dense open substack of $\mathscr{H}_{\varphi,N,\preceq\mu}$ and projects onto $\mathscr{D}_{\varphi,N,\mu}$. The morphism $\mathscr{H}_{\varphi,N,\mu} \to \mathscr{D}_{\varphi,N,\mu}$ has a section and is relatively representable by a vector bundle.*

(c) *The stack $\mathscr{H}_{\varphi,\mu}$ is a dense open substack of $\mathscr{H}_{\varphi,\preceq\mu}$ and projects onto $\mathscr{D}_{\varphi,\mu}$. The morphism of stacks $\mathscr{H}_{\varphi,\mu} \to \mathscr{D}_{\varphi,\mu}$ has a section and is relatively representable by a vector bundle.*

(d) *The stacks $\mathscr{H}_{\varphi,\preceq\mu}$, $\mathscr{H}_{\varphi,\mu}$ are irreducible of dimension $\sum_{\psi,j}(d+1-2j)\mu_{\psi,j}$, and $\mathscr{D}_{\varphi,\mu}$ is irreducible of dimension $\sum_\psi \#\{(i,j) : \mu_{\psi,i} > \mu_{\psi,j}\}$. The stacks $\mathscr{H}_{\varphi,\mu}$ and $\mathscr{D}_{\varphi,\mu}$ are smooth over $E_\mu$.*

(e) *The stacks $\mathscr{H}_{\varphi,N,\preceq\mu}$, $\mathscr{H}_{\varphi,N,\mu}$ are equidimensional of dimension $\sum_{\psi,j}(d+1-2j)\mu_{\psi,j}$, and $\mathscr{D}_{\varphi,N,\mu}$ is equidimensional of dimension $\sum_\psi \#\{(i,j) : \mu_{\psi,i} > \mu_{\psi,j}\}$. The stacks $\mathscr{H}_{\varphi,N,\mu}$ and $\mathscr{D}_{\varphi,N,\mu}$ are reduced, Cohen–Macaulay and generically smooth over $E_\mu$. The irreducible components of $\mathscr{H}_{\varphi,N,\preceq\mu}$, $\mathscr{H}_{\varphi,N,\mu}$ and $\mathscr{D}_{\varphi,N,\mu}$ are indexed by the possible Jordan types of the nilpotent endomorphism $N$.*

*Proof.* (a) The stacks are quotients of noetherian schemes of finite type over $E_\mu$ by the action of the smooth group scheme $(\mathrm{Res}_{K_0/\mathbb{Q}_p}\mathrm{GL}_{d,K_0})_{E_\mu}$ and hence are noetherian Artin stacks of finite type by [Laumon and Moret-Bailly 2000, 4.6.1, 4.7.1, 4.14].

(b) and (c) follow from the corresponding statements for $Q_{K,d,\mu}$ in Proposition 3.5.

(d) The covering spaces

$$\mathrm{Res}_{K_0/\mathbb{Q}_p}\mathrm{GL}_{d,K_0}\times_{\mathrm{Spec}\,\mathbb{Q}_p} Q_{K,d,\preceq\mu},$$
$$\mathrm{Res}_{K_0/\mathbb{Q}_p}\mathrm{GL}_{d,K_0}\times_{\mathrm{Spec}\,\mathbb{Q}_p} Q_{K,d,\mu},$$
$$\mathrm{Res}_{K_0/\mathbb{Q}_p}\mathrm{GL}_{d,K_0}\times_{\mathrm{Spec}\,\mathbb{Q}_p} \mathrm{Flag}_{K,d,\mu}$$

of these stacks are irreducible because $\mathrm{Res}_{K_0/\mathbb{Q}_p}\mathrm{GL}_{d,K_0}$ is geometrically irreducible. This implies the irreducibility of the stacks. The formulas for the dimension follow from [Laumon and Moret-Bailly 2000, pp. 98f] and Proposition 3.5, respectively the well known dimension formula for partial flag varieties. The smoothness follows from the smoothness of $\mathrm{Res}_{K_0/\mathbb{Q}_p}\mathrm{GL}_{d,K_0}\times_{\mathrm{Spec}\,\mathbb{Q}_p} Q_{K,d,\mu}$ and $\mathrm{Res}_{K_0/\mathbb{Q}_p}\mathrm{GL}_{d,K_0}\times_{\mathrm{Spec}\,\mathbb{Q}_p}\mathrm{Flag}_{K,d,\mu}$ by [Laumon and Moret-Bailly 2000, 4.14].

(e) As in (d) these results are direct consequences of the corresponding results on the covering spaces, which follow from Theorem 3.2. We only need to convince ourselves that the action of $(\mathrm{Res}_{K_0/\mathbb{Q}_p}\mathrm{GL}_{d,K_0})_{E_\mu}$ does not identify irreducible components of $P_{K_0,d}$. However this follows from the fact that the Jordan canonical forms of the nilpotent endomorphism $N$ at two distinct generic points $y_1$ and $y_2$ of $P_{K_0,d}$ are distinct by the description in Theorem 3.2. □

**Remark 3.7.** These stacks are not separated. Namely, let $\underline{D}$, $\underline{D}'$ be two $(\varphi, N)$-modules with Hodge–Pink lattice (respectively two $K$-filtered $(\varphi, N)$-modules) over $R$. Then $\mathrm{Isom}(\underline{D}, \underline{D}')$ is representable by an algebraic space, separated and of finite type over $R$; see [Laumon and Moret-Bailly 2000, Lemme 4.2]. The above stacks are separated over $E_\mu$ if and only if all these algebraic spaces $\mathrm{Isom}(\underline{D}, \underline{D}')$ are proper. This is not the case in general. For example let $R$ be a discrete valuation ring with fraction field $L$, let $D = D' = R\otimes_{\mathbb{Q}_p} K_0^d$ with $\Phi = \mathrm{id}$ and $N = 0$. Then every element $f \in L$ is an automorphism of $D\otimes_R L$, compatible with $\Phi$ and $N$. However, it extends to an automorphism of $D$ only if $f \in R^\times$.

## 4. Vector bundles on the open unit disc

Kisin [2006] related $K$-filtered $(\varphi, N)$-modules over $\mathbb{Q}_p$ to vector bundles on the open unit disc. This was generalized in [Hellmann 2013, §5] to families of $K$-filtered $\varphi$-modules with Hodge–Tate weights 0 and $-1$. In this section we generalize it to arbitrary families of $(\varphi, N)$-modules with Hodge–Pink lattice. For this purpose we work in the category $\mathrm{Ad}^{\mathrm{lft}}_{\mathbb{Q}_p}$ of adic spaces locally of finite type over $\mathrm{Spa}(\mathbb{Q}_p, \mathbb{Z}_p)$; see [Huber 1993; 1994; 1996; Hellmann 2013, §2.2]. Since the stacks $\mathscr{D}_{\varphi,\mu}$, $\mathscr{D}_{\varphi,N,\mu}$, $\mathscr{H}_{\varphi,\preceq\mu}$, $\mathscr{H}_{\varphi,N,\preceq\mu}$, $\mathscr{H}_{\varphi,\mu}$ and $\mathscr{H}_{\varphi,N,\mu}$ are quotients of quasiprojective schemes over $E_\mu$ they give rise to stacks on $\mathrm{Ad}^{\mathrm{lft}}_{E_\mu}$ which we denote by $\mathscr{H}^{\mathrm{ad}}_{\varphi,N,\mu}$, etc.

For $0 \le r < 1$ we write $\mathbb{B}_{[0,r]}$ for the closed disc of radius $r$ over $K_0$ in the category of adic spaces and denote by

$$\mathbb{U} = \varinjlim_{r\to 1} \mathbb{B}_{[0,r]}$$

the open unit disc. This is an open subspace of the closed unit disc (which is *not* identified with the set of all points $x$ in the closed unit disc with $|x| < 1$ in the adic setting). In the following we will always

write $u$ for the coordinate function on $\mathbb{B}_{[0,r]}$ and $\mathbb{U}$, i.e., we view

$$\mathbb{B}^{[0,r]} := \Gamma(\mathbb{B}_{[0,r]}, \mathcal{O}_{\mathbb{B}_{[0,r]}})$$

and $\mathbb{B}^{[0,1)} := \Gamma(\mathbb{U}, \mathcal{O}_{\mathbb{U}})$ as subrings of $K_0[\![u]\!]$.

Let $X \in \mathrm{Ad}^{\mathrm{lft}}_{\mathbb{Q}_p}$ be an adic space over $\mathbb{Q}_p$ and write

$$\mathscr{B}_X^{[0,r]} = \mathcal{O}_X \widehat{\otimes}_{\mathbb{Q}_p} \mathbb{B}^{[0,r]} = \mathrm{pr}_{X,*}\, \mathcal{O}_{X\times \mathbb{B}_{[0,r]}}$$
$$\mathscr{B}_X^{[0,1)} = \mathcal{O}_X \widehat{\otimes}_{\mathbb{Q}_p} \mathbb{B}^{[0,1)} = \mathrm{pr}_{X,*}\, \mathcal{O}_{X\times \mathbb{U}}$$

for the sheafified versions of the rings $\mathbb{B}^{[0,r]}$ and $\mathbb{B}^{[0,1)}$ where $\mathrm{pr}_X$ is the projection onto $X$. These are sheaves of topological $\mathcal{O}_X$-algebras on $X$.

We introduce the function

$$\lambda := \prod_{n\geq 0} \varphi^n(E(u)/E(0)) \in \mathbb{B}^{[0,1)}. \tag{4-1}$$

and the differential operator $N_\nabla := -u\lambda(d/du) : \mathbb{B}_{[0,1)} \to \mathbb{B}_{[0,1)}$. For any adic space $X \in \mathrm{Ad}^{\mathrm{lft}}_{\mathbb{Q}_p}$ we view $\lambda$ as a section of $\mathscr{B}_X^{[0,1)}$ and $N_\nabla$ as a differential operator on $\mathscr{B}_X^{[0,1)}$. The Frobenius $\varphi$ on $\mathcal{O}_X \otimes_{\mathbb{Q}_p} K_0[u]$ extends to a Frobenius endomorphism of $\mathscr{B}_X^{[0,1)}$ again denoted by $\varphi$ by means of $\varphi(u) = u^p$. These operators satisfy the relation

$$N_\nabla \varphi = p\, \frac{E(u)}{E(0)} \cdot \varphi\, N_\nabla. \tag{4-2}$$

**Definition 4.1.** A $(\varphi, N_\nabla)$-*module* $(\mathcal{M}, \Phi_\mathcal{M}, N_\nabla^\mathcal{M})$ over an adic space $X \in \mathrm{Ad}^{\mathrm{lft}}_{\mathbb{Q}_p}$ consists of a locally free sheaf $\mathcal{M}$ of finite rank on $X \times_{\mathbb{Q}_p} \mathbb{U}$, a differential operator $N_\nabla^\mathcal{M} : \mathcal{M} \to \mathcal{M}[1/\lambda]$ over $N_\nabla$, that is $N_\nabla^\mathcal{M}(fm) = -u\lambda(df/du)\cdot m + f\cdot N_\nabla^\mathcal{M}(m)$ for all sections $f$ of $\mathcal{O}_{X\times_{\mathbb{Q}_p}\mathbb{U}}$ and $m$ of $\mathcal{M}$, and an $\mathcal{O}_{X\times_{\mathbb{Q}_p}\mathbb{U}}$-linear isomorphism $\Phi_\mathcal{M} : (\varphi^*\mathcal{M})[1/E(u)] \xrightarrow{\sim} \mathcal{M}[1/E(u)]$, satisfying

$$N_\nabla^\mathcal{M} \circ \Phi_\mathcal{M} \circ \varphi = p\, \frac{E(u)}{E(0)} \cdot \Phi_\mathcal{M} \circ \varphi \circ N_\nabla^\mathcal{M}.$$

A *morphism* $\alpha : (\mathcal{M}, \Phi_\mathcal{M}, N_\nabla^\mathcal{M}) \to (\mathcal{N}, \Phi_\mathcal{N}, N_\nabla^\mathcal{N})$ between $(\varphi, N_\nabla)$-modules over $X$ is a morphism $\alpha : \mathcal{M} \to \mathcal{N}$ of sheaves satisfying $\alpha \circ \Phi_\mathcal{M} = \Phi_\mathcal{N} \circ \varphi^*(\alpha)$ and $N_\nabla^\mathcal{N} \circ \alpha = \alpha \circ N_\nabla^\mathcal{M}$.

**Remark 4.2.** (1) Note that it is not clear whether a $(\varphi, N_\nabla)$-module $\mathcal{M}$ is locally on $X$ free over $X \times \mathbb{U}$ and hence it is not clear whether $\mathrm{pr}_{X,*}\,\mathcal{M}$ is locally on $X$ a free $\mathscr{B}_X^{[0,1)}$-module. However it follows from [Kedlaya et al. 2014, Proposition 2.1.15] that $\mathrm{pr}_{X,*}\,\mathcal{M}$ is a finitely presented $\mathscr{B}_X^{[0,1)}$-module.

(2) The differential operator $N_\nabla^\mathcal{M}$ can be equivalently described as a connection

$$\nabla_\mathcal{M} : \mathcal{M} \to \mathcal{M} \otimes u^{-1}\Omega^1_{X\times\mathbb{U}/X}[1/\lambda]$$

when we set $\nabla_\mathcal{M}(m) := -(1/\lambda)N_\nabla^\mathcal{M}(m) \otimes du/u$. Then $N_\nabla^\mathcal{M}$ is recovered as the composition of $\nabla_\mathcal{M}$ followed by the map $u^{-1}\Omega^1_{X\times\mathbb{U}/X}[1/\lambda] \to \mathcal{M}$, $du \mapsto -u\lambda$.

Let $X \in \mathrm{Ad}^{\mathrm{lft}}_{\mathbb{Q}_p}$ be an adic space. We will show that the category of $(\varphi, N_\nabla)$-modules over $X$ is equivalent to the category of $(\varphi, N)$-modules with Hodge–Pink lattice over $X$ by defining two mutually quasi-inverse functors $\underline{\mathcal{M}}$ and $\underline{D}$.

To define $\underline{\mathcal{M}}$ let $\underline{D} = (D, \Phi, N, \mathfrak{q})$ be a $(\varphi, N)$-module with Hodge–Pink lattice over $X$. We denote by $\mathrm{pr} : X \times_{\mathbb{Q}_p} \mathbb{U} \to X \times_{\mathbb{Q}_p} K_0$ the projection and set $(\mathcal{D}, \Phi_{\mathcal{D}}) := \mathrm{pr}^*(D, \Phi)$. Then

$$\mathrm{pr}_{X,*}(\mathcal{D}, \Phi_{\mathcal{D}}) = (D, \Phi) \otimes_{(\mathcal{O}_X \otimes K_0)} \mathscr{B}_X^{[0,1)}.$$

We choose a $\mathbb{B}^+_{\mathcal{O}_X}$-automorphism $\eta_{\underline{D}}$ of $\mathfrak{p} := D \otimes_{\mathcal{O}_X \otimes K_0} \mathbb{B}^+_{\mathcal{O}_X}$ and we let $\iota_0 : \mathcal{D} \hookrightarrow D \otimes_{\mathcal{O}_X \otimes K_0} \mathbb{B}^+_{\mathcal{O}_X}$ be the embedding obtained as the composition of the natural inclusion $D \otimes_{\mathcal{O}_X \otimes K_0} \mathscr{B}_X^{[0,1)} \hookrightarrow D \otimes_{\mathcal{O}_X \otimes K_0} \mathbb{B}^+_{\mathcal{O}_X}$ composed with the automorphism $\eta_{\underline{D}}$. Here we follow Kisin [2006, §1.2] and choose

$$\eta_{\underline{D}} : D \otimes_{\mathcal{O}_X \otimes K_0} \mathbb{B}^+_{\mathcal{O}_X} \xrightarrow{\sim} D \otimes_{\mathcal{O}_X \otimes K_0} \mathbb{B}^+_{\mathcal{O}_X}, \quad d_0 \otimes f \mapsto \sum_i N^i(d_0) \otimes \frac{(-1)^i}{i!} \log\Big(1 - \frac{E(u)}{E(0)}\Big)^i \cdot f. \quad (4\text{-}3)$$

**Remark 4.3.** (1) Actually, Kisin introduces a formal variable $\ell_u$ over $\mathscr{B}_X^{[0,1)}$ which formally acts like $\log u$. He extends $\varphi$ to $\mathscr{B}_X^{[0,1)}[\ell_u]$ via $\varphi(\ell_u) = p\,\ell_u$, extends $N_\nabla$ to a derivation on $\mathscr{B}_X^{[0,1)}[\ell_u]$ via $N_\nabla(\ell_u) = -\lambda$, and defines $N$ as the $\mathscr{B}_X^{[0,1)}$-linear derivation on $\mathscr{B}_X^{[0,1)}[\ell_u]$ that acts as the differentiation of the formal variable $\ell_u$. Under the $\Phi$-equivariant identification

$$D[\ell_u]^{N=0} := \Big\{ \sum_{i=0}^{<\infty} d_i \ell_u^i : d_i \in D \text{ with } N\Big(\sum_i d_i \ell_u^i\Big) = 0 \Big\} \xleftarrow{\cong} D, \qquad \sum_{i=0}^{<\infty} \frac{(-1)^i}{i!} N^i(d_0) \ell_u^i \longleftarrow d_0,$$

Kisin's map $\iota_0 : D[\ell_u]^{N=0} \otimes_{\mathcal{O}_X \otimes K_0} \mathscr{B}_X^{[0,1)} \hookrightarrow \mathfrak{p}$, $\sum_i d_i \ell_u^i \otimes f \mapsto \sum_i d_i \otimes f \cdot (\log(u/\pi))^i$ corresponds to our $\iota_0$, because we identify $E(u)/E(0)$ with $1 - (u/\pi)$.

(2) Instead of the above $\eta_{\underline{D}}$ one could also choose $\eta_{\underline{D}} = \mathrm{id}_{\mathfrak{p}}$. This would lead to a few changes which we will comment on in Remark 4.11. Note that our $\eta_{\underline{D}}$ from (4-3) is different from $\mathrm{id}_{\mathfrak{p}}$ if $N \neq 0$.

For all $n \geq 0$ we now consider the map

$$\mathrm{pr}_{X,*}\,\mathcal{D}\Big[\frac{1}{\lambda}\Big] \xrightarrow{\Phi_{\mathcal{D}}^{-j}} \mathrm{pr}_{X,*}\,\varphi^{j*}\Big(\mathcal{D}\Big[\frac{1}{\lambda}\Big]\Big) = \mathrm{pr}_{X,*}\,\mathcal{D}\Big[\frac{1}{\lambda}\Big] \otimes_{\mathscr{B}_X^{[0,1)},\,\varphi^j} \mathscr{B}_X^{[0,1)} \xrightarrow{\varphi^{j*}\iota_0} \mathfrak{p}\Big[\frac{1}{E(u)}\Big] \otimes_{\mathbb{B}^+_{\mathcal{O}_X}} \varphi^j(\mathbb{B}^+_{\mathcal{O}_X}),$$

where we write $\varphi^{j*}\iota_0$ for $\iota_0 \otimes \mathrm{id}$. We set

$$\mathrm{pr}_{X,*}\,\mathcal{M} := \Big\{ m \in \mathrm{pr}_{X,*}\,\mathcal{D}\Big[\frac{1}{\lambda}\Big] : \varphi^{j*}\iota_0 \circ \Phi_{\mathcal{D}}^{-j}(m) \in \mathfrak{q} \otimes_{\mathbb{B}^+_{\mathcal{O}_X}} \varphi^j(\mathbb{B}^+_{\mathcal{O}_X}) \text{ for all } j \geq 0 \Big\}. \quad (4\text{-}4)$$

and we let $\mathcal{M}$ be the induced sheaf on $X \times_{\mathbb{Q}_p} \mathbb{U}$. Since $\lambda = (E(u)/E(0))\varphi(\lambda)$ the isomorphism $\Phi_{\mathcal{D}}$ induces an isomorphism $\Phi_{\mathcal{M}} : (\varphi^*\mathcal{M})[1/E(u)] \xrightarrow{\sim} \mathcal{M}[1/E(u)]$.

We want to show that $\mathcal{D}$ and $\mathcal{M}$ are locally free sheaves of finite rank on $X \times_{\mathbb{Q}_p} \mathbb{U}$. For $\mathcal{D}$ this follows from $\mathcal{D}|_{X \times \mathbb{B}_{[0,r]}} = D \otimes_{(\mathcal{O}_X \otimes K_0)} \mathcal{O}_{X \times \mathbb{B}_{[0,r]}}$. We work on a covering of $X$ by affinoids $Y = \mathrm{Spa}(A, A^+)$. Let

$h \in \mathbb{Z}$ be such that $\mathfrak{q} \subset E(u)^{-h}\mathfrak{p}$ on $Y$ and let $n$ be maximal such that $\varphi^n(E(u))$ is not a unit in $\mathbb{B}_{[0,r]}$, that is, such that $r^{p^n} \geq |\pi|$. Then $\mathcal{M}|_{Y\times\mathbb{B}_{[0,r]}}$ is defined by the exact sequence

$$0 \to \mathcal{M}|_{Y\times\mathbb{B}_{[0,r]}} \to \lambda^{-h}\mathcal{D}|_{Y\times\mathbb{B}_{[0,r]}} \xrightarrow{\bigoplus_{j=0}^{n} \varphi^{j*}\iota_0\circ\Phi_{\mathcal{D}}^{-j}} \bigoplus_{j=0}^{n} (E(u)^{-h}\mathfrak{p}/\mathfrak{q}) \otimes_{\mathbb{B}_A^+,\varphi^j} \varphi^j(\mathbb{B}_A^+) \to 0.$$

The $A \otimes_{\mathbb{Q}_p} K_0[u]$-module $E(u)^{-h}\mathfrak{p}/\mathfrak{q}$ is locally free over $A$, say of rank $k$. The endomorphism $\varphi : K_0[u] \to K_0[u]$ makes the target $K_0[u]$ into a free module of rank $p$ over the source $K_0[u]$. Therefore $(E(u)^{-h}\mathfrak{p}/\mathfrak{q}) \otimes_{\mathbb{B}_A^+,\varphi^j} \varphi^j(\mathbb{B}_A^+)$ is locally free over $A$ of rank $p^j k$. Since the affinoid algebra $A$ is noetherian and $\mathbb{B}_{[0,r]}$ is a principal ideal domain by [Lazard 1962, Corollary of Proposition 4] also $\Gamma(Y \times \mathbb{B}_{[0,r]}, \mathcal{O}_{Y\times\mathbb{B}_{[0,r]}}) = A\widehat{\otimes}_{\mathbb{Q}_p}\mathbb{B}_{[0,r]}$ is noetherian. So $\Gamma(Y \times \mathbb{B}_{[0,r]}, \mathcal{M})$ is finitely generated over $A\widehat{\otimes}_{\mathbb{Q}_p}\mathbb{B}_{[0,r]}$ and flat over $A$. The residue field of each maximal ideal $\mathfrak{m} \subset A\widehat{\otimes}_{\mathbb{Q}_p}\mathbb{B}_{[0,r]}$ is finite over $\mathbb{Q}_p$ by [Bosch et al. 1984, Corollary 6.1.2/3]. Therefore $\mathfrak{n} = \mathfrak{m} \cap A$ is a maximal ideal of $A$. By the elementary divisor theorem $A/\mathfrak{n} \otimes_A \Gamma(Y \times \mathbb{B}_{[0,r]}, \mathcal{M})$ is free over the product of principal ideal domains $A/\mathfrak{n} \otimes_{\mathbb{Q}_p} \mathbb{B}_{[0,r]}$. Therefore $\Gamma(Y \times \mathbb{B}_{[0,r]}, \mathcal{M})$ is locally free of rank $d$ over $A\widehat{\otimes}_{\mathbb{Q}_p}\mathbb{B}_{[0,r]}$ by [EGA IV$_3$ 1966, Theorem 11.3.10]. This shows that $\mathcal{M}$ is a locally free sheaf of rank $d$ on $X \times_{\mathbb{Q}_p} \mathbb{U}$.

We equip $\mathcal{M}$ with a differential operator $N_\nabla^{\mathcal{M}}$ over $N_\nabla$. On $\lambda^{-h}\mathcal{D} = D \otimes_{(\mathcal{O}_X\otimes K_0)} \lambda^{-h}\mathscr{B}_X^{[0,1)}$ we have the differential operator $N_\nabla^{\mathcal{D}} := N \otimes \lambda + \mathrm{id}_D \otimes N_\nabla$

$$\lambda^{-h}\mathcal{D} \xrightarrow{N\otimes\lambda+\mathrm{id}_D\otimes N_\nabla} \lambda^{-h}\mathcal{D}, \quad d \otimes \lambda^{-h} f \mapsto N(d) \otimes \lambda^{1-h} f + d \otimes \left(hu\lambda^{-h} f \frac{d\lambda}{du} - u\lambda^{1-h}\frac{df}{du}\right) \quad (4\text{-}5)$$

with $d \in D$ and $f \in \mathscr{B}_X^{[0,1)}$. Its image lies in $\lambda^{-h}\mathcal{D}$. If $E(u)^n\mathfrak{p} \subset \mathfrak{q} \subset E(u)^{-h}\mathfrak{p}$ then $\lambda^n\mathcal{D} \subset \mathcal{M} \subset \lambda^{-h}\mathcal{D}$. Thus $N_\nabla^{\mathcal{D}}(\mathcal{M}) \subset \lambda^{-h}\mathcal{D} \subset \lambda^{-h-n}\mathcal{M}$ and we let $N_\nabla^{\mathcal{M}}$ be the restriction of $N_\nabla^{\mathcal{D}}$ to $\mathcal{M}$. The equation $N_\nabla^{\mathcal{M}} \circ \Phi_{\mathcal{M}} \circ \varphi = p(E(u)/E(0)) \cdot \Phi_{\mathcal{M}} \circ \varphi \circ N_\nabla^{\mathcal{M}}$ is satisfied because it is satisfied on $\mathcal{D}$ by (4-2). Therefore we have constructed a $(\varphi, N_\nabla)$-module $\underline{\mathcal{M}}(D, \Phi, N, \mathfrak{q}) := (\mathcal{M}, \Phi_{\mathcal{M}}, N_\nabla^{\mathcal{M}})$ over $X$. Note that in terms of Kisin's description of $\mathcal{D} \cong D[\ell_u]^{N=0} \otimes_{(\mathcal{O}_X\otimes K_0)} \mathscr{B}_X^{[0,1)}$ the differential operator $N_\nabla^{\mathcal{M}}$ is given as $\mathrm{id}_D \otimes N_\nabla$.

**Example 4.4.** The $(\varphi, N)$-modules with Hodge–Pink lattice from Example 2.9, corresponding to the cyclotomic character, give rise to the following $(\varphi, N_\nabla)$-modules of rank 1 over $X = \mathrm{Spa}(\mathbb{Q}_p, \mathbb{Z}_p)$. For $\underline{D} = (K_0, \Phi = p^{-1}, N = 0, \mathfrak{q} = E(u)\mathfrak{p})$ we obtain $\mathcal{D} = (\mathscr{B}_X^{[0,1)}, \Phi_{\mathcal{D}} = p^{-1}, N_\nabla)$ and $\mathcal{M} = \lambda\mathscr{B}_X^{[0,1)}$. On the basis vector $\lambda$ of $\mathcal{M}$ the actions of $\Phi_{\mathcal{D}}$ and $N_\nabla$ are given by $\Phi_{\mathcal{D}}(\varphi(\lambda)) = p^{-1}\varphi(\lambda) = (E(0)/pE(u))\,\lambda$ and $N_\nabla(\lambda) = -u(d\lambda/du)\,\lambda$. So we find $\underline{\mathcal{M}}(\underline{D}) \cong (\mathscr{B}_X^{[0,1)}, \Phi_{\mathcal{M}} = (E(0)/pE(u)), N_\nabla^{\mathcal{M}})$ with $N_\nabla^{\mathcal{M}}(f) = N_\nabla(f) - u(d\lambda/du)\,f$. Similarly for $\underline{D} = (K_0, \Phi = p, N = 0, \mathfrak{q} = E(u)^{-1}\mathfrak{p})$ we obtain $\mathcal{D} = (\mathscr{B}_X^{[0,1)}, \Phi_{\mathcal{D}} = p, N_\nabla)$ and $\mathcal{M} = \lambda^{-1}\mathscr{B}_X^{[0,1)}$ which leads to $\underline{\mathcal{M}}(\underline{D}) \cong (\mathscr{B}_X^{[0,1)}, \Phi_{\mathcal{M}} = (pE(u)/E(0)), N_\nabla^{\mathcal{M}})$ with $N_\nabla^{\mathcal{M}}(f) = N_\nabla(f) + u(d\lambda/du)\,f$.

To define the quasi-inverse functor $\underline{D}$ let $(\mathcal{M}, \Phi_{\mathcal{M}}, N_\nabla^{\mathcal{M}})$ be a $(\varphi, N_\nabla)$-module over $X$. We denote by $e : X \times_{\mathbb{Q}_p} K_0 \to X \times_{\mathbb{Q}_p} \mathbb{U}$ the isomorphism $x \mapsto (x, 0)$ onto the closed subspace defined by $u = 0$. Let $(D, \Phi, N) := e^*(\mathcal{M}, \Phi_{\mathcal{M}}, N_\nabla^{\mathcal{M}})$. It is a $(\varphi, N)$-module over $X$ because $N$ is clearly $\mathcal{O}_X \otimes_{\mathbb{Q}_p} K_0$-linear

and $e^*(E(u)/E(0)) = 1$ implies $N \circ \Phi = p \cdot \Phi \circ \varphi^* N$. By [Pappas and Rapoport 2009, Proposition 5.2] there is a unique $\mathcal{O}_{X \times_{\mathbb{Q}_p} \mathbb{U}}$-linear isomorphism

$$\xi : \mathrm{pr}^* D\Big[\frac{1}{\lambda}\Big] \xrightarrow{\sim} \mathcal{M}\Big[\frac{1}{\lambda}\Big] \tag{4-6}$$

satisfying $\xi \circ \mathrm{pr}^* \Phi = \Phi_{\mathcal{M}} \circ \varphi^* \xi$ and $e^* \xi = \mathrm{id}_D$. In particular the composition $\mathrm{pr}^* \Phi \circ (\varphi^* \xi)^{-1} = \xi^{-1} \circ \Phi_{\mathcal{M}}$ induces an isomorphism $\varphi^* \mathcal{M} \otimes \mathbb{B}^+_{\mathcal{O}_X} \xrightarrow{\sim} D \otimes_{(\mathcal{O}_X \otimes K_0)} \mathbb{B}^+_{\mathcal{O}_X} = \mathfrak{p}$ of $\mathbb{B}^+_{\mathcal{O}_X}$-modules. We set $\mathfrak{q} := \eta_{\underline{D}} \circ (\xi \otimes \mathrm{id}_{\mathbb{B}_{\mathcal{O}_X}})^{-1}(\mathcal{M} \otimes \mathbb{B}^+_{\mathcal{O}_X})$. Then $\underline{D}(\mathcal{M}, \Phi_{\mathcal{M}}, N^{\mathcal{M}}_{\nabla}) := (D, \Phi, N, \mathfrak{q})$ is a $(\varphi, N)$-module with Hodge–Pink lattice over $X$ by Lemma 2.7 and the following lemma.

**Lemma 4.5.** *Locally on a covering of $X$ by affinoids $Y = \mathrm{Spa}(A, A^+)$ there exist integers $h, n$ with $E(u)^n \Phi_{\mathcal{M}}(\varphi^* \mathcal{M}) \subset \mathcal{M} \subset E(u)^{-h} \Phi_{\mathcal{M}}(\varphi^* \mathcal{M})$ such that the quotients*

$$E(u)^{-h} \Phi_{\mathcal{M}}(\varphi^* \mathcal{M})/\mathcal{M} \quad \text{and} \quad \mathcal{M}/E(u)^n \Phi_{\mathcal{M}}(\varphi^* \mathcal{M})$$

*are finite locally free over $A$.*

*Proof.* We may assume that $X = Y = \mathrm{Spa}(A, A^+)$ is affinoid. Then the existence of $h$ and $n$ follows from the finiteness of $\mathcal{M}$ and $\varphi^* \mathcal{M}$. Let $\mathfrak{m} \subset A$ be a maximal ideal and set $L = A/\mathfrak{m}$. Let $|\pi| < r < 1$ and set $\widetilde{\mathcal{M}} := \Gamma(Y \times_{\mathbb{Q}_p} \mathbb{B}_{[0,r]}, \mathcal{M})$ and $\widetilde{\varphi^* \mathcal{M}} := \Gamma(Y \times_{\mathbb{Q}_p} \mathbb{B}_{[0,r]}, \varphi^* \mathcal{M})$. Then $\mathcal{M}/E(u)^n \Phi_{\mathcal{M}}(\varphi^* \mathcal{M}) \cong \widetilde{\mathcal{M}} E(u)^n \Phi_{\mathcal{M}}(\widetilde{\varphi^* \mathcal{M}})$. Consider the exact sequence

$$0 \to E(u)^n \widetilde{\varphi^* \mathcal{M}} \xrightarrow{\Phi_{\mathcal{M}}} \widetilde{\mathcal{M}} \to \mathcal{M}/E(u)^n \Phi_{\mathcal{M}}(\varphi^* \mathcal{M}) \to 0$$

in which the first map is injective because $E(u)$ is a nonzero-divisor in $A \widehat{\otimes}_{\mathbb{Q}_p} \mathbb{B}_{[0,r]}$. We tensor the sequence with $L$ over $A$ to obtain the exact sequence of $L \otimes_{\mathbb{Q}_p} \mathbb{B}_{[0,r]}$-modules

$$0 \to T \to L \otimes_A E(u)^n \widetilde{\varphi^* \mathcal{M}} \xrightarrow{\mathrm{id}_L \otimes \Phi_{\mathcal{M}}} L \otimes_A \widetilde{\mathcal{M}} \to L \otimes_A \big(\mathcal{M}/E(u)^n \Phi_{\mathcal{M}}(\varphi^* \mathcal{M})\big) \to 0$$

with $T = \mathrm{Tor}_1^A\big(L,\ \mathcal{M}/E(u)^n \Phi_{\mathcal{M}}(\varphi^* \mathcal{M})\big)$. Since $L \otimes_{\mathbb{Q}_p} K_0$ is a product of fields, $L \otimes_{\mathbb{Q}_p} \mathbb{B}_{[0,r]}$ is a product of principal ideal domains by [Lazard 1962, Corollary of Proposition 4]. Since $E(u)^{n+h}$ annihilates $L \otimes_A \mathcal{M}/E(u)^n \Phi_{\mathcal{M}}(\varphi^* \mathcal{M})$ the latter is a torsion module over $L \otimes_{\mathbb{Q}_p} \mathbb{B}_{[0,r]}$. It follows that $\mathrm{id}_L \otimes \Phi_{\mathcal{M}}$ is a morphism of free modules of the same rank over a product of principal ideal domains whose cokernel is a torsion module. It is a direct consequence of the classification of finitely generated modules over a principal ideal domain that the map $\mathrm{id}_L \otimes \Phi_{\mathcal{M}}$ then has to be injective. It follows that

$$0 = T = \mathrm{Tor}_1^A\big(L,\ \mathcal{M}/E(u)^n \Phi_{\mathcal{M}}(\varphi^* \mathcal{M})\big) = \mathrm{Tor}_1^{A_{\mathfrak{m}}}\big((A_{\mathfrak{m}}/\mathfrak{m}A_{\mathfrak{m}}),\ (\mathcal{M}/E(u)^n \Phi_{\mathcal{M}}(\varphi^* \mathcal{M}))_{\mathfrak{m}}\big).$$

Since $(\mathcal{M}/E(u)^n \Phi_{\mathcal{M}}(\varphi^* \mathcal{M}))_{\mathfrak{m}}$ is finite over the noetherian local ring $A_{\mathfrak{m}}$ it is locally free by the local criterion of flatness [Eisenbud 1995, Theorem 6.8]. It follows that $\mathcal{M}/E(u)^n \Phi_{\mathcal{M}}(\varphi^* \mathcal{M})$ is locally free as an $A$-module. Finally, the two last objects in the short exact sequence

$$0 \to E(u)^{-h} \Phi_{\mathcal{M}}(\varphi^* \mathcal{M})/\mathcal{M} \xrightarrow{\cdot E(u)^{n+h}} \mathcal{M}/E(u)^{n+h} \mathcal{M} \to \mathcal{M}/E(u)^n \Phi_{\mathcal{M}}(\varphi^* \mathcal{M}) \to 0$$

are flat and hence so is the first (all its higher Tor-terms have to vanish). As $E(u)^{-h} \Phi_{\mathcal{M}}(\varphi^* \mathcal{M})/\mathcal{M}$ is also finite as an $A$-module it follows that it is finite and locally free over $A$. □

**Theorem 4.6.** *For every adic space $X \in \mathrm{Ad}^{\mathrm{lft}}_{\mathbb{Q}_p}$ the functors $\underline{\mathcal{M}}$ and $\underline{D}$ constructed above are mutually quasi-inverse equivalences between the category of $(\varphi, N)$-modules with Hodge–Pink lattice over $X$ and the category of $(\varphi, N_\nabla)$-modules over $X$.*

*Proof.* We must show that the functors are mutually quasi-inverse. To prove one direction let $(D, \Phi, N, \mathfrak{q})$ be a $(\varphi, N)$-module with Hodge–Pink lattice over $X$ and let $(\mathcal{M}, \Phi_{\mathcal{M}}, N_\nabla^{\mathcal{M}}) = \underline{\mathcal{M}}(D, \Phi, N, \mathfrak{q})$. By construction $e^*\mathcal{M} = D$, and under this equality $e^*\Phi_{\mathcal{M}}$ corresponds to $\Phi$. Since $e^*\lambda = 1$, formula (4-5) shows that $e^*N_\nabla^{\mathcal{M}}$ corresponds to $N$ on $D$. By the uniqueness of the map $\xi$ from (4-6), its inverse $\xi^{-1}$ equals the inclusion $\mathcal{M} \hookrightarrow \mathcal{D}[1/\lambda]$, by which we defined $\mathcal{M}$. This shows that $\eta_{\underline{D}} \circ (\xi \otimes \mathrm{id}_{\mathbb{B}_{\mathcal{O}_X}})^{-1}(\mathcal{M} \otimes \mathbb{B}^+_{\mathcal{O}_X})$ equals $\mathfrak{q}$ and that $\underline{D} \circ \underline{\mathcal{M}} = \mathrm{id}$.

Conversely let $(\mathcal{M}, \Phi_{\mathcal{M}}, N_\nabla^{\mathcal{M}})$ be a $(\varphi, N_\nabla)$-module over $X$ and let $(D, \Phi, N, \mathfrak{q}) = \underline{D}(\mathcal{M}, \Phi_{\mathcal{M}}, N_\nabla^{\mathcal{M}})$. Via the isomorphism $\xi$ from (4-6), $\mathcal{M}$ is a $\varphi$-submodule of $\mathrm{pr}^*D[1/\lambda]$. By construction of $\mathfrak{q}$ and

$$\underline{\mathcal{M}}(\underline{D}(\mathcal{M}, \Phi_{\mathcal{M}}, N_\nabla^{\mathcal{M}})) \subset \mathrm{pr}^*D[1/\lambda],$$

the latter submodule coincides with $\mathcal{M}$ modulo all powers of $E(u)$. Since both submodules have a Frobenius which is an isomorphism outside $\mathrm{V}(E(u))$ they are equal on all of $X \times_{\mathbb{Q}_p} \mathbb{U}$. It remains to show that $N_\nabla^{\mathcal{M}}$ is compatible with $N_\nabla^{\mathrm{pr}^*D}$ under the isomorphism $\xi : \mathrm{pr}^*D[1/\lambda] \xrightarrow{\sim} \mathcal{M}[1/\lambda]$. We follow [Kisin 2006, Lemma 1.2.12(3)] and let $\sigma := \xi \circ N_\nabla^{\mathrm{pr}^*D} - N_\nabla^{\mathcal{M}} \circ \xi$. Then $\sigma : \mathrm{pr}^*D[1/\lambda] \to \mathcal{M}[1/\lambda]$ is $\mathcal{O}_{X \times_{\mathbb{Q}_p} \mathbb{U}}$-linear and it suffices to show that $\sigma(D) = 0$. By (4-5) both $N_\nabla^{\mathrm{pr}^*D}$ and $N_\nabla^{\mathcal{M}}$ reduce to $N$ modulo $u$. Therefore $\sigma(D) \subset u\mathcal{M}[1/\lambda]$. One checks that $\sigma \circ \Phi_{\mathrm{pr}^*D} \circ \varphi = p(E(u)/E(0)) \cdot \Phi_{\mathcal{M}} \circ \varphi \circ \sigma$ and this implies

$$\sigma(D) = \sigma \circ \Phi_{\mathrm{pr}^*D}(\varphi^*D) = p\frac{E(u)}{E(0)} \cdot \Phi_{\mathcal{M}} \circ \varphi^*\Big(u\mathcal{M}\Big[\frac{1}{\lambda}\Big]\Big) \subset u^p\mathcal{M}\Big[\frac{1}{\lambda}\Big].$$

By induction $\sigma(D) \subset u^{p^i}\mathcal{M}[1/\lambda]$ for all $i$ and hence $\sigma(D) = 0$. This shows that also $\underline{\mathcal{M}} \circ \underline{D}$ is isomorphic to the identity and proves the theorem. □

**Corollary 4.7.** *The stack $\mathscr{H}^{\mathrm{ad}}_{\varphi,N,\preceq\mu}$ is isomorphic to the stack whose groupoid of $X$-valued points for $X \in \mathrm{Ad}^{\mathrm{lft}}_{E_\mu}$ consists of $(\varphi, N_\nabla)$-modules $(\mathcal{M}, \Phi_{\mathcal{M}}, N_\nabla^{\mathcal{M}})$ over $X$ satisfying*

$$\textstyle\bigwedge^j_{\mathcal{O}_{X\times\mathbb{U}}}\mathcal{M} \subset E(u)^{-\mu_1-\cdots-\mu_j} \cdot \bigwedge^j_{\mathcal{O}_{X\times\mathbb{U}}}\Phi_{\mathcal{M}}(\varphi^*\mathcal{M})$$

*with equality for $j = \mathrm{rk}\,\mathcal{M}$. Here $\mu_i$ is viewed as a $\mathbb{Z}$-valued function on $X \times_{\mathbb{Q}_p} K$.*

*Proof.* This follows from the definition of the functor $\underline{D}$, in particular the definition of the Hodge–Pink lattice. □

**Definition 4.8.** We define substacks $\mathscr{H}^{\nabla}_{\varphi,N,\mu} \subset \mathscr{H}^{\mathrm{ad}}_{\varphi,N,\mu}$, $\mathscr{H}^{\nabla}_{\varphi,N,\preceq\mu} \subset \mathscr{H}^{\mathrm{ad}}_{\varphi,N,\preceq\mu}$, $\mathscr{H}^{\nabla}_{\varphi,\mu} \subset \mathscr{H}^{\mathrm{ad}}_{\varphi,\mu}$ and $\mathscr{H}^{\nabla}_{\varphi,\preceq\mu} \subset \mathscr{H}^{\mathrm{ad}}_{\varphi,\preceq\mu}$. For an adic space $X \in \mathrm{Ad}^{\mathrm{lft}}_{E_\mu}$ the groupoid $\mathscr{H}^{\nabla}_{\varphi,N,\mu}(X)$ consists of those $(D, \varphi, N, \mathfrak{q}) \in \mathscr{H}^{\mathrm{ad}}_{\varphi,N,\mu}(X)$ for which the associated $(\varphi, N_\nabla)$-module $(\mathcal{M}, \Phi_{\mathcal{M}}, N_\nabla^{\mathcal{M}})$ satisfies $N_\nabla^{\mathcal{M}}(\mathcal{M}) \subset \mathcal{M}$. The groupoids $\mathscr{H}^{\nabla}_{\varphi,N,\preceq\mu}(X)$, $\mathscr{H}^{\nabla}_{\varphi,\mu}(X)$ and $\mathscr{H}^{\nabla}_{\varphi,\preceq\mu}(X)$ are defined by the same condition. (Note that on the latter two $N = 0$, but $N_\nabla^{\mathcal{M}} \neq 0$.)

**Theorem 4.9.** *The substacks $\mathscr{H}^{\nabla}_{\varphi,N,\mu} \subset \mathscr{H}^{\mathrm{ad}}_{\varphi,N,\mu}$, $\mathscr{H}^{\nabla}_{\varphi,N,\preceq\mu} \subset \mathscr{H}^{\mathrm{ad}}_{\varphi,N,\preceq\mu}$, $\mathscr{H}^{\nabla}_{\varphi,\mu} \subset \mathscr{H}^{\mathrm{ad}}_{\varphi,\mu}$ and $\mathscr{H}^{\nabla}_{\varphi,\preceq\mu} \subset \mathscr{H}^{\mathrm{ad}}_{\varphi,\preceq\mu}$ are Zariski closed substacks. The substack $\mathscr{H}^{\nabla}_{\varphi,N,\mu}$ coincides with the image of the zero section of the vector bundle $\mathscr{H}^{\mathrm{ad}}_{\varphi,N,\mu} \to \mathscr{D}^{\mathrm{ad}}_{\varphi,N,\mu}$.*

**Remark 4.10.** We can consider a family of $(\varphi, N_\nabla)$-modules over $\mathscr{D}^{\mathrm{ad}}_{\varphi,N,\mu}$. We pull back the canonical family of $(\varphi, N_\nabla)$-modules on $\mathscr{H}^{\mathrm{ad}}_{\varphi,N,\mu}$ along the zero section. Then for $x \in \mathscr{D}^{\mathrm{ad}}_{\varphi,N,\mu}(\mathbb{Q}_p)$ the fiber of this family at $x$ coincides with the $(\varphi, N_\nabla)$-module that Kisin [2006] associates with the filtered $(\varphi, N)$-module defined by $x$.

**Remark 4.11.** If instead of the isomorphism $\eta_{\underline{D}}$ from (4-3) we choose $\eta_{\underline{D}} = \mathrm{id}_{\mathfrak{p}}$ as in Remark 4.3 (2), the above results remain valid, except that $\mathscr{H}^{\nabla}_{\varphi,N,\mu}$ coincides with the image of a different section. This section is obtained by composing the zero section with the inverse

$$\eta_{\underline{D}}^{-1} : d \otimes f \mapsto \textstyle\sum_i N^i(d) \otimes (1/i!) \log\bigl(1 - (E(u)/E(0))\bigr)^i \cdot f$$

of the automorphism $\eta_{\underline{D}}$. It sends a filtration $\mathcal{F}^\bullet$ to $\eta_{\underline{D}}^{-1}\bigl(\sum_{i\in\mathbb{Z}} E(u)^{-i}(\mathcal{F}^i D_K) \otimes_{R\otimes K} \mathbb{B}_R^+\bigr)$. Note that both sections coincide on the closed substack $\mathscr{D}^{\mathrm{ad}}_{\varphi,\mu}$ where $N = 0$.

*Proof of Theorem 4.9.* To prove that the substacks are closed let $\underline{D} \in \mathscr{H}^{\mathrm{ad}}_{\varphi,N,\preceq\mu}(X)$ for an adic space $X \in \mathrm{Ad}^{\mathrm{lft}}_{E_\mu}$ and let $(M, \Phi_{\mathcal{M}}, N_\nabla^{\mathcal{M}}) = \underline{\mathcal{M}}(\underline{D})$ be the associated $(\varphi, N_\nabla)$-module over $X$. Locally on $X$ there is an integer $h$ with $N_\nabla^{\mathcal{M}}(\mathcal{M}) \subset \lambda^{-h}\mathcal{M}$ by Lemma 2.7 and the construction of $N_\nabla^{\mathcal{M}}$. The quotients $\bigl(\lambda^{-h}\mathcal{M}/\mathcal{M}\bigr) \otimes \bigl(\mathscr{B}_X^{[0,1)}/(\varphi^n(E(u))^h)\bigr)$ are finite locally free as $\mathcal{O}_X$-modules for all $n \geq 0$. Now the condition $N_\nabla^{\mathcal{M}}(\mathcal{M}) \subset \mathcal{M}$ is equivalent to the vanishing of the images under $N_\nabla^{\mathcal{M}}$ of a set of generators of $\mathcal{M}$ in $\bigl(\lambda^{-h}\mathcal{M}/\mathcal{M}\bigr) \otimes \bigl(\mathscr{B}_X^{[0,1)}/(\varphi^n(E(u))^h)\bigr)$ for each $n \geq 0$. Due to [EGA I 1971, Lemma 9.7.9.1] the latter is represented by a Zariski closed subspace of $X$.

We show that the closed substack $\mathscr{H}^{\nabla}_{\varphi,N,\mu}$ of $\mathscr{H}^{\mathrm{ad}}_{\varphi,N,\mu}$ coincides with the image of the zero section. Since $N_\nabla^{\mathcal{D}}$ on $\mathcal{D} := D \otimes_{\mathcal{O}_X\otimes K_0} \mathscr{B}^{[0,1)}_{\mathcal{O}_X}$ induces the differential operator $\mathrm{id}_D \otimes N_\nabla$ on $\mathfrak{p} := D \otimes_{\mathcal{O}_X\otimes K_0} \mathbb{B}^+_{\mathcal{O}_X}$ under the map $i_0 = \eta_{\underline{D}} \circ \text{inclusion} : \mathcal{D} \hookrightarrow \mathfrak{p}$ from (4-3), it follows directly that the image of the zero section is contained in $\mathscr{H}^{\nabla}_{\varphi,N,\mu}$. To prove the converse we may work on the coverings $X := (P_{K_0,d} \times_{\mathbb{Q}_p} Q_{K,d,\mu})^{\mathrm{ad}}$ of $\mathscr{H}^{\mathrm{ad}}_{\varphi,N,\mu}$ and $(P_{K_0,d} \times_{\mathbb{Q}_p} \mathrm{Flag}_{K,d,\mu})^{\mathrm{ad}}$ of $\mathscr{D}^{\mathrm{ad}}_{\varphi,N,\mu}$ because the zero section and $\mathscr{H}^{\nabla}_{\varphi,N,\mu}$ are both invariant under the action of $(\mathrm{Res}_{K_0/\mathbb{Q}_p} \mathrm{GL}_{d,K_0})_{E_\mu}$. We first claim that both have the same underlying topological space. By [Bosch et al. 1984, Corollary 6.1.2/3] this can be checked on $L$-valued points of $X$ for finite extensions $L$ of $E_\mu$. For those it was proved by Kisin [2006, Lemma 1.2.12(4)] that the universal Hodge–Pink lattice $\mathfrak{q}$ at $L$ lies in the image of the zero section if the pullback $\underline{M}$ to $L$ of the universal $(\varphi, N_\nabla)$-module on $X$ has holomorphic $N_\nabla^{\mathcal{M}}$. From this our claim follows.

To prove equality as closed subspaces of $X$ we look at a closed point $x \in X$ and its complete local ring $\widehat{\mathcal{O}}_{X,x}$. Let $\mathfrak{m}_x \subset \widehat{\mathcal{O}}_{X,x}$ be the maximal ideal, let $I \subset \widehat{\mathcal{O}}_{X,x}$ be the ideal defining $\mathscr{H}^{\nabla}_{\varphi,N,\mu}$, and set $R_n := \widehat{\mathcal{O}}_{X,x}/(\mathfrak{m}_x^n + I)$. Then $R_n$ is a finite-dimensional $\mathbb{Q}_p$-vector space by [Bosch et al. 1984, Corollary 6.1.2/3]. We consider the universal $\underline{D}_{R_n} = (D, \Phi, N, \mathfrak{q})$ over $R_n$ by restriction of scalars from $R_n$ to $\mathbb{Q}_p$ as a $(\varphi, N)$-module $\widetilde{\underline{D}}$ with Hodge–Pink lattice over $\mathbb{Q}_p$ of rank $(\dim_{\mathbb{Q}_p} R_n)(\mathrm{rk}_{R_n\otimes K_0} D) = \dim_{K_0} D$. It is equipped with a ring homomorphism $R_n \to \mathrm{End}(\widetilde{\underline{D}})$. Since $N_\nabla^{\mathcal{M}}$ is holomorphic on $\underline{\mathcal{M}}(\underline{D}_{R_n})$, Kisin [2006,

Lemma 1.2.12(4)] tells us again that $\mathfrak{q} = \mathfrak{q}(\mathcal{F}^\bullet)$ for the filtration $\mathcal{F}^\bullet = \mathcal{F}^\bullet_{\mathfrak{q}}$ from Remark 2.8. This shows that the ideal $J$ defining the zero section in $X$ vanishes in $R_n$. Since this holds for all $n$, the ideals $I$ and $J$ are equal in $\widehat{\mathcal{O}}_{X,x}$. As $x$ was arbitrary, they coincide on all of $X$ and this proves the theorem. □

## 5. Weak admissibility

Similar to the case of filtrations, one can define a notion of weak admissibility for $(\varphi, N)$-modules with Hodge–Pink lattice and develop a Harder–Narasimhan formalism. Compare also [Hellmann 2011, §2] for the following. Recall that $f = [K_0 : \mathbb{Q}_p]$ and $e = [K : K_0]$.

**Definition 5.1.** Let $L$ be a field with a valuation $v_L : L \to \Gamma_L \cup \{0\}$ in the sense of [Huber 1993, §2, Definition] and set $\Gamma_L^{\mathbb{Q}} := \Gamma_L \otimes_{\mathbb{Z}} \mathbb{Q}$.

(i) Let $\underline{D} = (D, \Phi, N)$ be a $(\varphi, N)$-module over $L$. Then define

$$t_N(\underline{D}) := v_L(\det_L \Phi^f)^{1/f^2} \in \Gamma_L^{\mathbb{Q}}.$$

If $L \supset K_0$ we are in the situation of Remark 2.4 and have $t_N(\underline{D}) = v_L(\det_L(\Phi^f)_0)^{1/f}$.

(ii) Let $\underline{D} = (D, \Phi, N, \mathcal{F}^\bullet)$ be a $K$-filtered $(\varphi, N)$-module over $L$. Then

$$t_H(\underline{D}) := \frac{1}{ef} \sum_{i \in \mathbb{Z}} i \dim_L(\mathcal{F}^i D_K / \mathcal{F}^{i+1} D_K) \in \mathbb{Q}.$$

(iii) Let $\underline{D} = (D, \Phi, N, \mathfrak{q})$ be a $(\varphi, N)$-module with Hodge–Pink lattice of rank $d$ over $L$. Then we set

$$t_H(\underline{D}) := \frac{1}{ef}\big(\dim_L(\mathfrak{q}/t^n\mathfrak{p}) - \dim_L(\mathfrak{p}/t^n\mathfrak{p})\big) = \frac{1}{ef}\dim_L(\mathfrak{q}/t^n\mathfrak{p}) \; - \; n \operatorname{rk} \underline{D} \in \mathbb{Q}$$

for $n \gg 0$, which is independent of $n$ whenever $t^n\mathfrak{p} \subset \mathfrak{q}$. If $L$ is an extension of $\widetilde{K}$ and $(\mu_\psi)_\psi = \mu_{\underline{D}}(\operatorname{Spec} L)$ is the Hodge polygon of $\underline{D}$ (see Definition 2.11) then $t_H(\underline{D}) := (1/(ef)) \sum_\psi \mu_{\psi,1} + \cdots + \mu_{\psi,d}$. If the $\psi$-component $\mathfrak{q}_\psi$ satisfies $\bigwedge^d \mathfrak{q}_\psi = t^{-h_\psi} \bigwedge^d \mathfrak{p}_\psi$ then $t_H(\underline{D}) = (1/(ef)) \sum_\psi h_\psi$. Moreover $t_H(\underline{D}) = t_H(D, \Phi, N, \mathcal{F}^\bullet_{\mathfrak{q}})$.

(iv) Let $\underline{D}$ be a $(\varphi, N)$-module with Hodge–Pink lattice (or a $K$-filtered $(\varphi, N)$-module) over $L$. Then its *slope* is defined to be

$$\lambda(\underline{D}) := (v_L(p)^{t_H(\underline{D})} \cdot t_N(\underline{D})^{-1})^{1/d} \in \Gamma_L^{\mathbb{Q}}.$$

**Definition 5.2.** (i) A $(\varphi, N)$-module with Hodge–Pink lattice $\underline{D} = (D, \Phi, N, \mathfrak{q})$ over a field $L$ endowed with a valuation is called *semistable* if $\lambda(\underline{D}') \geq \lambda(\underline{D})$ for all $\underline{D}' = \big(D', \Phi|_{\varphi^* D'}, N|_{D'}, \mathfrak{q} \cap D' \otimes_{L \otimes_{\mathbb{Q}_p} K_0} \mathbb{B}_L\big)$ where $D' \subset D$ is a free $L \otimes_{\mathbb{Q}_p} K_0$-submodule stable under $\Phi$ and $N$.

(ii) A $K$-filtered $(\varphi, N)$-module $\underline{D} = (D, \Phi, N, \mathcal{F}^\bullet)$ over $L$ is called *semistable* if $\lambda(\underline{D}') \geq \lambda(\underline{D})$ for all $\underline{D}' = \big(D', \Phi|_{\varphi^* D'}, N|_{D'}, \mathcal{F}^\bullet \cap D'_K\big)$ where $D' \subset D$ is a free $L \otimes_{\mathbb{Q}_p} K_0$-submodule stable under $\Phi$ and $N$.

(iii) A $(\varphi, N)$-module with Hodge–Pink lattice (or a $K$-filtered $(\varphi, N)$-module) is called *weakly admissible* if it is semistable of slope 1.

**Lemma 5.3.** *Let $(D, \Phi, N, \mathcal{F}^\bullet)$ be a $K$-filtered $(\varphi, N)$-module over a valued field $L$ and let $(D, \Phi, N, \mathfrak{q})$ denote the $(\varphi, N)$-module with Hodge–Pink lattice associated to $(D, \Phi, N, \mathcal{F}^\bullet)$ by the zero section $\mathcal{F}^\bullet \mapsto \mathfrak{q} = \mathfrak{q}(\mathcal{F}^\bullet)$ of Remark 2.8 (3). Then $(D, \Phi, N, \mathcal{F}^\bullet)$ is weakly admissible if and only if $(D, \Phi, N, \mathfrak{q})$ is.*

*Proof.* It is obvious from the definitions that $t_H(D, \Phi, N, \mathcal{F}^\bullet) = t_H(D, \Phi, N, \mathfrak{q}(\mathcal{F}^\bullet))$. Further we have to test on the same subobjects $D' \subset D$. Hence the claim follows from the fact

$$\mathfrak{q}(\mathcal{F}^\bullet \cap D'_K) = \mathfrak{q}(\mathcal{F}^\bullet) \cap D' \otimes_{L \otimes_{\mathbb{Q}_p} K_0} \mathbb{B}_L,$$

which is obvious from the description of $\mathfrak{q}(-)$ in Remark 2.8 (3) by choosing an $L \otimes_{\mathbb{Q}_p} K$-basis of $D_K$ adapted to the submodules $\mathcal{F}^i D'_K$ and $\mathcal{F}^i D_K$. □

**Proposition 5.4.** *Let $(D, \Phi, N, \mathfrak{q})$ be a $(\varphi, N)$-module with Hodge–Pink lattice defined over some valued field $L$. Then there is a unique Harder–Narasimhan filtration*

$$0 = D_0 \subset D_1 \subset \cdots \subset D_r = D$$

*of $(D, \Phi, N, \mathfrak{q})$, by free $L \otimes_{\mathbb{Q}_p} K_0$-submodules stable under $\Phi$ and $N$ such that the subquotients $D_i / D_{i-1}$ with their induced Hodge–Pink lattice are semistable of slope $\lambda_i \in \Gamma_L \otimes \mathbb{Q}$ and $\lambda_1 < \lambda_2 < \cdots < \lambda_r$.*

*Proof.* This is the usual Harder–Narasimhan formalism; see [Fargues and Fontaine 2018, 5.5.1] for a fairly general exposition. See also [Hellmann 2011, Proposition 2.19]. □

**Corollary 5.5.** *Let $(D, \Phi, N, \mathfrak{q})$ be a $(\varphi, N)$-module with Hodge–Pink lattice over $L$ and let $L'$ be an extension of $L$ with valuation $v_{L'}$ extending the valuation $v_L$. Then $(D, \Phi, N, \mathfrak{q})$ is weakly admissible if and only if $(D', \Phi', N', \mathfrak{q}') = (D \otimes_L L', \Phi \otimes \mathrm{id}, N \otimes \mathrm{id}, \mathfrak{q} \otimes_L L')$ is weakly admissible.*

*Proof.* This is similar to [Hellmann 2011, Corollary 2.22].

If $(D', \Phi', N', \mathfrak{q}')$ is weakly admissible, then every $(\Phi, N)$-stable subobject $D_1 \subset D$ defines a $(\Phi', N')$-stable subobject $D'_1 = D_1 \otimes_L L'$ of $D'$ such that

$$\lambda(D_1, \Phi|_{\varphi^* D_1}, N|_{D_1}, \mathfrak{q} \cap D_1 \otimes_{L \otimes K_0} \mathbb{B}_L) = \lambda(D'_1, \Phi'|_{\varphi^* D'_1}, N'|_{D'_1}, \mathfrak{q}' \cap D'_1 \otimes_{L' \otimes K_0} \mathbb{B}_{L'}).$$

It follows that $(D, \Phi, N, \mathfrak{q})$ is weakly admissible, as $(D', \Phi', N', \mathfrak{q}')$ is.

Now assume that $(D, \Phi, N, \mathfrak{q})$ is weakly admissible. We may reduce to the case where $L'$ is finitely generated over $L$ and $\mathrm{Aut}(L'/L)$ is large enough. As in the proof of [Hellmann 2011, Corollary 2.22] one shows that the action of $\mathrm{Aut}(L'/L)$ preserves the slope of $\Phi'$-stable subobjects of $D'$. Hence the Harder–Narasimhan filtration of $D'$ descends to $D$. As $D$ is weakly admissible, this filtration can only have one step. □

**Theorem 5.6.** *Let $\mu$ be a cocharacter as in* (2-5) *with reflex field $E_\mu$. Then the groupoid*

$$X \mapsto \left\{ (D, \Phi, N, \mathfrak{q}) \in \mathscr{H}_{\varphi, N, \preceq \mu}(X) \mid D \otimes \kappa(x) \text{ is weakly admissible for all } x \in X \right\}$$

*is an open substack $\mathscr{H}^{\mathrm{ad,wa}}_{\varphi, N, \preceq \mu}$ of $\mathscr{H}^{\mathrm{ad}}_{\varphi, N, \preceq \mu}$ on the category of adic spaces locally of finite type over $E_\mu$.*

*Proof.* This is similar to the proof of [Hellmann 2013, Theorem 4.1].

It follows from Corollary 5.5 that $\mathscr{H}^{\mathrm{ad,wa}}_{\varphi,N,\preceq\mu}$ is indeed a stack, i.e., weak admissibility may be checked over an fpqc-covering. Hence it suffices to show that the weakly admissible locus is open in

$$X_\mu := P_{K_0,d} \times_{\mathbb{Q}_p} Q_{K,d,\preceq\mu}.$$

Let us denote by $Z_i$ the projective $P_{K_0,d}$-scheme whose $S$-valued points are given by pairs $(x, U)$ with $x = (g, N) \in P_{K_0,d}(S) \subset (\mathrm{Res}_{K_0/\mathbb{Q}_p} \mathrm{GL}_{d,K_0}) \times (\mathrm{Res}_{K_0/\mathbb{Q}_p} \mathrm{Mat}_{d\times d})$ and an $\mathcal{O}_S \otimes K_0$-subspace $U \subset \mathcal{O}_S \otimes K_0^{\oplus d}$ which is locally on $S$ free of rank $i$, a direct summand as $\mathcal{O}_S$-module, and stable under the action of $\Phi_g = g \cdot \varphi$ and $N$. This is a closed subscheme of the product $P_{K_0,d} \times_{\mathbb{Q}_p} \mathrm{Quot}_{\mathcal{O}^d \mid K_0 \mid \mathbb{Q}_p}$ (where $\mathrm{Quot}_{\mathcal{O}^d \mid K_0 \mid \mathbb{Q}_p}$ is Grothendieck's Quot-scheme which is projective over $\mathbb{Q}_p$; see [Grothendieck 1962, n°221, Theorem 3.1] or [Altman and Kleiman 1980, Theorem 2.6]), cut out by the invariance conditions under $\Phi_g$ and $N$. Further write $f_i \in \Gamma(Z_i, \mathcal{O}_{Z_i})$ for the global section defined by

$$f_i(g, U) = \det(g \cdot \varphi)^f|_U = \det\big(g \cdot \varphi(g) \cdots \varphi^{f-1}(g)\big)^f|_U,$$

where $f = [K_0 : \mathbb{Q}_p]$, and where the determinant is the determinant as $\mathcal{O}_{Z_i}$-modules. Write $U$ for the pullback of the universal $(\Phi, N)$-invariant subspace on $Z_i$ to the product $Z_i \times Q_{K,d,\preceq\mu}$, write $\mathfrak{q}$ for the pullback of the universal $\mathbb{B}^+$-lattice on $Q_{K,d,\preceq\mu}$ to $Z_i \times Q_{K,d,\preceq\mu}$, and write $\mathfrak{p} = (\mathbb{B}^+)^{\oplus d}$ for the pullback of the tautological $\mathbb{B}^+$-lattice $D \otimes \mathbb{B}^+$ on $P_{K_0,d}$ to $Z_i \times Q_{K,d,\preceq\mu}$. Fix integers $n, h$ with $t^n\mathfrak{p} \subset \mathfrak{q} \subset t^{-h}\mathfrak{p}$ and consider the complex of finite locally free sheaves on $Z_i \times Q_{K,d,\preceq\mu}$

$$P_\bullet:\ P_1 := t^{-h}\mathfrak{p}/t^n\mathfrak{p} \xrightarrow{\delta} t^{-h}\mathfrak{p}/\mathfrak{q} \oplus \big(D/U \otimes t^{-h}\mathbb{B}^+/t^n\mathbb{B}^+\big) =: P_0$$

given by the canonical projection $D \twoheadrightarrow D/U$ in the second summand. Let $T_1$ be the functor from the category of quasicoherent sheaves on $Z_i \times Q_{K,d,\preceq\mu}$ to itself defined by

$$T_1 : M \mapsto \ker\big(\delta \otimes \mathrm{id}_M : P_1 \otimes M \to P_0 \otimes M\big).$$

If $M = \kappa(y)$ for a point $y = (g_y, N_y, U_y, \mathfrak{q}_y) \in Z_i \times Q_{K,d,\preceq\mu}$ then $T_1(\kappa(y)) = (\mathfrak{q}_y \cap \mathfrak{p}_{i,y}[1/t])/t^n\mathfrak{p}_{i,y}$, where we write $\mathfrak{p}_{i,y} := U_y \otimes_{\kappa(y)\otimes K_0} \mathbb{B}^+_{\kappa(y)}$. We consider the function

$$h_i : Z_i \times Q_{K,d,\preceq\mu} \to \mathbb{Q},$$

$$y \mapsto \frac{1}{ef} \dim_{\kappa(y)} T_1(\kappa(y)) - ni = t_H\Big(U_y, g_y(\mathrm{id} \otimes \varphi)|_{U_y}, N|_{U_y}, \mathfrak{q}_y \cap \mathfrak{p}_{i,y}\Big[\frac{1}{t}\Big]\Big). \tag{5-1}$$

We write $Z_i^{\mathrm{ad}}$ and $Q^{\mathrm{ad}}_{K,d,\preceq\mu}$ for the adic spaces associated to the varieties $Z_i$ and $Q_{K,d,\preceq\mu}$. Similarly we write $h_i^{\mathrm{ad}}$ for the function on the adic spaces $Z_i^{\mathrm{ad}} \times Q^{\mathrm{ad}}_{K,d,\preceq\mu}$ defined by the same formula as in (5-1). By semicontinuity [EGA III$_2$ 1963, Théorème 7.6.9], the sets

$$Y_{i,m} = \{y \in Z_i^{\mathrm{ad}} \times Q^{\mathrm{ad}}_{K,d,\preceq\mu} \mid h_i(y) \geq m\}$$

are closed and hence proper over $X_\mu^{\mathrm{ad}} = P^{\mathrm{ad}}_{K_0,d} \times_{\mathbb{Q}_p} Q^{\mathrm{ad}}_{K,d,\preceq\mu}$. We write

$$\mathrm{pr}_{i,m} : Y_{i,m} \to X_\mu^{\mathrm{ad}}$$

for the canonical, proper projection.

If we write $X_0 \subset X_\mu^{\mathrm{ad}}$ for the open subset of all $(D, \Phi, N, \mathfrak{q})$ such that $\lambda(D, \Phi, N, \mathfrak{q}) = 1$, then

$$X_0 \smallsetminus X_\mu^{\mathrm{wa}} = X_0 \cap \bigcup_{i,m} \mathrm{pr}_{i,m}\left(\left\{y \in Y_{i,m} \mid v_y(f_i) > v_y(p)^{f^2 m}\right\}\right), \tag{5-2}$$

where the union runs over $1 \le i \le d-1$ and $m \in \mathbb{Z}$. Indeed, let $x = (D, \Phi, N, \mathfrak{q})$ be an $L$-valued point of $X_0$, then any proper $(\Phi, N)$-stable subspace of $D' \subset D$ defines (for some $1 \le i \le d-1$) a point $y = (D', \mathfrak{q})$ of $Z_i \times Q_{K,d,\preceq\mu}$ mapping to $x$. This subspace violates the weak admissibility condition if and only if

$$v_y(f_i) = t_N(D', \Phi|_{\varphi^* D'})^{f^2} > v_y(p)^{f^2 t_H(D', \Phi, \mathfrak{q} \cap (D' \otimes \mathbb{B}_{\kappa(x)}))} = v_y(p)^{f^2 h_i(y)},$$

and hence (5-2) follows. On the other hand the union

$$\bigcup_{i,m} \mathrm{pr}_{i,m}\left(\left\{y \in Y_{i,m} \mid v_y(f_i) > v_y(p)^{f^2 m}\right\}\right)$$

is a finite union, because $Y_{i,m} = \varnothing$ for $m > hi$ and $Y_{i,m} = Z_i^{\mathrm{ad}} \times Q_{K,d,\preceq\mu}^{\mathrm{ad}}$ for $m \le -ni$. Therefore the union is closed by properness of the map $\mathrm{pr}_{i,m}$ and the definition of the topology on an adic space. The theorem follows from this. □

We define the subgroupoid $\mathscr{H}_{\varphi,\preceq\mu}^{\mathrm{ad,wa}} \subset \mathscr{H}_{\varphi,\preceq\mu}^{\mathrm{ad}}$ as follows. Given an adic space $X$ and $(D, \Phi, N, \mathfrak{q}) \in \mathscr{H}_{\varphi,\preceq\mu}^{\mathrm{ad}}$, we say that $(D, \Phi, N, \mathfrak{q}) \in \mathscr{H}_{\varphi,\preceq\mu}^{\mathrm{ad,wa}}$ if and only if $(D, \Phi, N, \mathfrak{q}) \otimes \kappa(x)$ is weakly admissible for all points $x \in X$. We define the subgroupoids $\mathscr{H}_{\varphi,N,\mu}^{\mathrm{ad,wa}} \subset \mathscr{H}_{\varphi,N,\mu}^{\mathrm{ad}}$, $\mathscr{H}_{\varphi,\mu}^{\mathrm{ad,wa}} \subset \mathscr{H}_{\varphi,\mu}^{\mathrm{ad}}$, $\mathscr{D}_{\varphi,N,\mu}^{\mathrm{ad,wa}} \subset \mathscr{D}_{\varphi,N,\mu}^{\mathrm{ad}}$ and $\mathscr{D}_{\varphi,\mu}^{\mathrm{ad,wa}} \subset \mathscr{D}_{\varphi,\mu}^{\mathrm{ad}}$ in the same manner.

**Corollary 5.7.** *The subgroupoids $\mathscr{H}_{\varphi,\preceq\mu}^{\mathrm{ad,wa}} \subset \mathscr{H}_{\varphi,\preceq\mu}^{\mathrm{ad}}$, $\mathscr{H}_{\varphi,N,\mu}^{\mathrm{ad,wa}} \subset \mathscr{H}_{\varphi,N,\mu}^{\mathrm{ad}}$, $\mathscr{H}_{\varphi,\mu}^{\mathrm{ad,wa}} \subset \mathscr{H}_{\varphi,\mu}^{\mathrm{ad}}$, $\mathscr{D}_{\varphi,N,\mu}^{\mathrm{ad,wa}} \subset \mathscr{D}_{\varphi,N,\mu}^{\mathrm{ad}}$ and $\mathscr{D}_{\varphi,\mu}^{\mathrm{ad,wa}} \subset \mathscr{D}_{\varphi,\mu}^{\mathrm{ad}}$ are open substacks.*

*Proof.* This follows by pulling back $\mathscr{H}_{\varphi,N,\preceq\mu}^{\mathrm{ad,wa}} \subset \mathscr{H}_{\varphi,N,\preceq\mu}^{\mathrm{ad}}$ along the morphisms $\mathscr{H}_{\varphi,\preceq\mu}^{\mathrm{ad}} \to \mathscr{H}_{\varphi,N,\preceq\mu}^{\mathrm{ad}}$, $\mathscr{H}_{\varphi,N,\mu}^{\mathrm{ad}} \to \mathscr{H}_{\varphi,N,\preceq\mu}^{\mathrm{ad}}$, $\mathscr{H}_{\varphi,\mu}^{\mathrm{ad}} \to \mathscr{H}_{\varphi,N,\preceq\mu}^{\mathrm{ad}}$, $\mathscr{D}_{\varphi,N,\mu}^{\mathrm{ad}} \to \mathscr{H}_{\varphi,N,\preceq\mu}^{\mathrm{ad}}$ and $\mathscr{D}_{\varphi,\mu}^{\mathrm{ad}} \to \mathscr{H}_{\varphi,N,\preceq\mu}^{\mathrm{ad}}$, respectively. Here we use the fact that the zero sections $\mathscr{D}_{\varphi,N,\mu}^{\mathrm{ad}} \to \mathscr{H}_{\varphi,N,\preceq\mu}^{\mathrm{ad}}$ and $\mathscr{D}_{\varphi,\mu}^{\mathrm{ad}} \to \mathscr{H}_{\varphi,N,\preceq\mu}^{\mathrm{ad}}$ preserve weak admissibility by Lemma 5.3. □

**Remark 5.8.** Note that the projection

$$\mathrm{pr} : \mathscr{H}_{\varphi,N,\mu}^{\mathrm{ad}} \to \mathscr{D}_{\varphi,N,\mu}^{\mathrm{ad}}$$

does not preserve weak admissibility. We always have $\mathrm{pr}^{-1}(\mathscr{D}_{\varphi,N,\mu}^{\mathrm{ad,wa}}) \subset \mathscr{H}_{\varphi,N,\mu}^{\mathrm{ad,wa}}$ and hence especially any section of the vector bundle $\mathscr{H}_{\varphi,N,\mu}^{\mathrm{ad}} \to \mathscr{D}_{\varphi,N,\mu}^{\mathrm{ad}}$ maps the weakly admissible locus to the weakly admissible locus.

Indeed, let $\underline{D} = (D, \Phi, N, \mathfrak{q})$ be a point of $\mathscr{H}_{\varphi,N,\mu}^{\mathrm{ad}}$ over a field $L$ whose image $(D, \Phi, N, \mathcal{F}_{\mathfrak{q}}^\bullet)$ in $\mathscr{D}_{\varphi,N,\mu}^{\mathrm{ad}}$ is weakly admissible. Let $D' \subset D$ be an $L \otimes_{\mathbb{Q}_p} K_0$-submodule which is stable under $\Phi$ and $N$, and set $\mathfrak{p}' := D' \otimes_{L \otimes K_0} \mathbb{B}_L^+$. Then $\mathfrak{q}' := \mathfrak{q} \cap \mathfrak{p}'[1/t]$ satisfies $t^i \mathfrak{q}' \cap \mathfrak{p}' \subset t^i \mathfrak{q} \cap \mathfrak{p}$ and $\mathcal{F}_{\mathfrak{q}'}^i D'_K \subset \mathcal{F}_{\mathfrak{q}}^i D_K \cap D'_K$.

This implies

$$t_H(D', \Phi|_{\varphi^* D'}, N|_{D'}, \mathfrak{q}') = t_H(D', \Phi|_{\varphi^* D'}, N|_{D'}, \mathcal{F}^\bullet_{\mathfrak{q}'}) \le t_H(D', \Phi|_{\varphi^* D'}, N|_{D'}, \mathcal{F}^\bullet_{\mathfrak{q}} \cap D'_K)$$

with equality for $D' = D$, and

$$v_L(p)^{t_H(D', \Phi|_{\varphi^* D'}, N|_{D'}, \mathfrak{q}')} \ge v_L(p)^{t_H(D', \Phi|_{\varphi^* D'}, N|_{D'}, \mathcal{F}^\bullet_{\mathfrak{q}} \cap D'_K)} \ge t_N(D', \Phi|_{\varphi^* D'})$$

with equality for $D' = D$, because $(D, \Phi, N, \mathcal{F}^\bullet_{\mathfrak{q}})$ is weakly admissible. Therefore also $\underline{D}$ is weakly admissible. This proves that $\mathrm{pr}^{-1}(\mathscr{D}^{\mathrm{ad,wa}}_{\varphi,N,\mu}) \subset \mathscr{H}^{\mathrm{ad,wa}}_{\varphi,N,\mu}$. However, in general this inclusion is strict as can be seen from the following example.

**Example 5.9.** Let $K = K_0 = \mathbb{Q}_p$, $d = 2$ and $\mu = (2, 0)$. We consider points $\underline{D} = (D, \Phi, N, \mathcal{F}^\bullet)$ in $\mathscr{D}^{\mathrm{ad}}_{\varphi,N,\mu}$ over a field $L$ with $\Phi = p\,\mathrm{Id}_2$ and $N = 0$. The filtration is of the form $D = \mathcal{F}^0 D \supset \mathcal{F}^1 D = \mathcal{F}^2 D = \binom{u}{v} \cdot L \supset \mathcal{F}^3 D = (0)$ for some $\binom{u}{v} \in D$. None of these points is weakly admissible, because the subspace $D' = \binom{u}{v} \cdot L \subset D$ has $t_N(\underline{D}') = v_L(p)$ and $\mathcal{F}^2 D' = D'$, whence $t_H(\underline{D}') = 2$ and $\lambda(\underline{D}') = v_L(p) < 1$.

The preimage of such a point in $\mathscr{H}^{\mathrm{ad}}_{\varphi,N,\mu}$ is given by a Hodge–Pink lattice $\mathfrak{q}$ with $\mathfrak{p} \subset \mathfrak{q} \subset t^{-2}\mathfrak{p}$ with Hodge weights 0 and $-2$. This means that $\mathfrak{q} = \mathfrak{p} + \binom{u+tu'}{v+tv'} \cdot t^{-2}\mathbb{B}^+_L$ for some $\binom{u'}{v'} \in D$. If the vectors $\binom{u}{v}$ and $\binom{u'}{v'}$ are linearly dependent over $L$ then $\underline{D} = (D, \Phi, N, \mathfrak{q})$ is not weakly admissible, because the subspace $D' = \binom{u}{v} \cdot L \subset D$ has $t_N(\underline{D}') = v_L(p)$ and $\mathfrak{q}' := \mathfrak{q} \cap D' \otimes_L \mathbb{B}_L = t^{-2} D' \otimes_L \mathbb{B}^+_L$, whence $t_H(\underline{D}') = 2$ and $\lambda(\underline{D}') = v_L(p) < 1$.

On the other hand, if the vectors $\binom{u}{v}$ and $\binom{u'}{v'}$ are linearly *independent* over $L$ then $\underline{D} = (D, \Phi, N, \mathfrak{q})$ is weakly admissible, because then $\mathfrak{q}' \subset t^{-1} D' \otimes_L \mathbb{B}^+_L$ for any subspace $D' = \binom{a}{b} \cdot L \subset D$, whence $t_N(\underline{D}') = v_L(p)$, $t_H(\underline{D}') \le 1$ and $\lambda(\underline{D}') \ge 1$. Indeed, $\binom{a}{b} \cdot t^{-2} \in \mathfrak{q}'$ would imply that $\binom{a}{b} \cdot t^{-2} \equiv \binom{u+tu'}{v+tv'} \cdot t^{-2} \cdot (c + tc') \equiv c\binom{u}{v} \cdot t^{-2} + \big(c'\binom{u}{v} + c\binom{u'}{v'}\big)t^{-1} \bmod \mathfrak{p}$ for $c, c' \in L$. This implies $\binom{a}{b} = c\binom{u}{v}$ and $c'\binom{u}{v} + c\binom{u'}{v'} = 0$ contradicting the linear independence.

Thus the weakly admissible locus $\mathscr{H}^{\mathrm{ad,wa}}_{\varphi,\mu}$ in the fiber of $\mathscr{H}_{\varphi,N,\mu}$ over the point $(\Phi, N) = (p\,\mathrm{Id}_2, 0)$ in $P_{\mathbb{Q}_p,2}$ equals the complement of the zero section, while this fiber in $\mathscr{D}^{\mathrm{ad,wa}}_{\varphi,N,\mu}$ is empty; see also Lemma 5.3.

We end this section by remarking that the weakly admissible locus is determined by the rigid analytic points, i.e., those points of an adic space whose residue field is a finite extension of $\mathbb{Q}_p$.

**Lemma 5.10.** *Let $X$ be an adic space locally of finite type over $E_\mu$ and let $f : X \to \mathscr{H}^{\mathrm{ad}}_{\varphi,N,\preceq\mu}$ be a morphism defined by a $(\varphi, N)$-module with Hodge–Pink lattice $\underline{D}$. Then $f$ factors over $\mathscr{H}^{\mathrm{ad,wa}}_{\varphi,N,\preceq\mu}$ if and only if $\underline{D} \otimes \kappa(x)$ is weakly admissible for all rigid analytic points $x \in X$.*

*Proof.* One implication is obvious and the other one is an easy application of the maximum modulus principle. It is proven along the same lines as [Hellmann 2013, Proposition 4.3]. □

**Remark 5.11.** The analogous statements for the stacks $\mathscr{H}^{\mathrm{ad,wa}}_{\varphi,\preceq\mu} \subset \mathscr{H}^{\mathrm{ad}}_{\varphi,\preceq\mu}$, $\mathscr{H}^{\mathrm{ad,wa}}_{\varphi,N,\mu} \subset \mathscr{H}^{\mathrm{ad}}_{\varphi,N,\mu}$, $\mathscr{H}^{\mathrm{ad,wa}}_{\varphi,\mu} \subset \mathscr{H}^{\mathrm{ad}}_{\varphi,\mu}$, $\mathscr{D}^{\mathrm{ad,wa}}_{\varphi,N,\mu} \subset \mathscr{D}^{\mathrm{ad}}_{\varphi,N,\mu}$ and $\mathscr{D}^{\mathrm{ad,wa}}_{\varphi,\mu} \subset \mathscr{D}^{\mathrm{ad}}_{\varphi,\mu}$ are also true and are a direct consequence of their construction.

## 6. The étale locus

Let us denote by $\mathbb{B}_{[r,s]}$ the closed annulus over $K_0$ of inner radius $r$ and outer radius $s$ for some $r, s \in [0,1) \cap p^{\mathbb{Q}}$. For an adic space $X \in \mathrm{Ad}^{\mathrm{lft}}_{\mathbb{Q}_p}$ we write

$$\mathscr{A}_X^{[0,1)} = \mathrm{pr}_{X,*}\,\mathcal{O}^+_{X\times\mathbb{U}} \subset \mathscr{B}_X^{[0,1)} = \mathrm{pr}_{X,*}\,\mathcal{O}_{X\times\mathbb{U}},$$
$$\mathscr{A}_X^{[r,s]} = \mathrm{pr}_{X,*}\,\mathcal{O}^+_{X\times\mathbb{B}_{[r,s]}} \subset \mathscr{B}_X^{[r,s]} = \mathrm{pr}_{X,*}\,\mathcal{O}_{X\times\mathbb{B}_{[r,s]}}.$$

The Frobenius $\varphi$ on $\mathscr{B}_X^{[0,1)}$ restricts to a ring homomorphism $\varphi$ on $\mathscr{A}_X^{[0,1)}$. For this section we adapt the notation from [Hellmann 2013] and write $r_i = r^{1/p^i}$. Then $\varphi$ restricts to a homomorphism

$$\varphi : \mathscr{B}_X^{[r,s]} \to \mathscr{B}_X^{[r_1,s_1]}.$$

**Definition 6.1.** A *$\varphi$-module of finite height* over $\mathscr{A}_X^{[0,1)}$ is an $\mathscr{A}_X^{[0,1)}$-module $\mathfrak{M}$ which is locally on $X$ free of finite rank over $\mathscr{A}_X^{[0,1)}$ together with an injective morphism $\Phi : \varphi^*\mathfrak{M} \to \mathfrak{M}$ of $\mathscr{A}_X^{[0,1)}$-modules such that $\operatorname{coker}\Phi$ is killed by some power of $E(u) \in W[\![u]\!] \subset \mathscr{A}_X^{[0,1)}$.

Inspired by Example 4.4 we define the $(\varphi, N_\nabla)$-module $\mathscr{B}_X^{[0,1)}(1)$ over $X$ to be

$$(\mathscr{B}_X^{[0,1)}, \Phi_{\mathcal{M}} = pE(u)/E(0), N_\nabla^{\mathcal{M}})$$

with $N_\nabla^{\mathcal{M}}(f) = N_\nabla(f) + u(d\lambda/du)\,f$. For an integer $n \in \mathbb{Z}$ we set

$$\mathscr{B}_X^{[0,1)}(n) := \mathscr{B}_X^{[0,1)}(1)^{\otimes n} = (\mathscr{B}_X^{[0,1)}, (pE(u)/E(0))^n, N_\nabla^{\mathcal{M}})$$

with $N_\nabla^{\mathcal{M}}(f) = N_\nabla(f) + nu(d\lambda/du)\,f$. Given a $(\varphi, N_\nabla)$-module $(\mathcal{M}, \Phi_{\mathcal{M}})$ on $X$ we write $(\mathcal{M}, \Phi_{\mathcal{M}})(n)$ for the twist $\mathcal{M} \otimes_{\mathscr{B}_X^{[0,1)}} \mathscr{B}_X^{[0,1)}(n)$. Note that $p/E(0) \in W^\times$ since $E(u)$ is an Eisenstein polynomial. Thus for $n \geq 0$ we have an obvious integral model $\mathscr{A}_X^{[0,1)}(n)$ for $\mathscr{B}^{[0,1)}(n)$ which is a $\varphi$-module of finite height over $\mathscr{A}_X^{[0,1)}$ (by forgetting the $N_\nabla$-action). Further we write $\mathbb{A}^{[0,1)}(n) = \mathscr{A}_{\mathrm{Spa}(\mathbb{Q}_p,\mathbb{Z}_p)}(n)$ for the $W[\![u]\!]$-module of rank 1 with basis $e$ on which $\Phi$ acts via $\Phi(e) = (E(u)/(pE(0)))^n e$.

**Definition 6.2.** Take $(\mathcal{M}, \Phi_{\mathcal{M}}, N_\nabla^{\mathcal{M}})$ to be a $(\varphi, N_\nabla)$-module over an adic space $X \in \mathrm{Ad}^{\mathrm{lft}}_{\mathbb{Q}_p}$.

(i) The module $\mathcal{M}$ is called *étale* if there exists an fpqc-covering $(U_i \to X)$, an integer $n \geq 0$ and $\varphi$-modules $(\mathfrak{M}_i, \Phi_{\mathfrak{M}_i})$ of finite height over $\mathscr{A}_{U_i}^{[0,1)}$ such that

$$(\mathcal{M}, \Phi_{\mathcal{M}})(n)|_{U_i} = (\mathfrak{M}_i, \Phi_{\mathfrak{M}_i}) \otimes_{\mathscr{A}_{U_i}^{[0,1)}} \mathscr{B}_{U_i}^{[0,1)}.$$

(ii) Let $x \in X$; then $\mathcal{M}$ is called *étale at $x$* if there exists an integer $n \geq 0$ and a $(\kappa(x)^+ \otimes_{\mathbb{Z}_p} W)[\![u]\!]$-lattice $\mathfrak{M} \subset \mathcal{M}(n) \otimes \kappa(x)$ such that

$$E(u)^h\mathfrak{M} \subset \Phi_{\mathcal{M}(n)}(\varphi^*\mathfrak{M}) \subset \mathfrak{M}$$

for some integer $h \geq 0$.

**Theorem 6.3.** *Let $X$ be an adic space locally of finite type over $\mathbb{Q}_p$ and let $(\mathcal{M}, \Phi)$ be a $(\varphi, N_\nabla)$-module. Then the subset*

$$X^{\mathrm{int}} = \{x \in X \mid \mathcal{M} \text{ is étale at } x\}$$

*is open and the restriction $\mathcal{M}|_{X^{\mathrm{int}}}$ is étale.*

This is similar to the proof of [Hellmann 2013, Theorem 7.6]. However, we need to make a few generalizations as we cannot rely on a reduced universal case. Given an affinoid algebra $A$ and $r, s \in [0, 1) \cap p^{\mathbb{Q}}$ we write

$$\mathbb{B}_A^{[r,s]} = \Gamma(\mathrm{Spa}(A, A^\circ), \mathscr{B}^{[r,s]}_{\mathrm{Spa}(A,A^\circ)}) = A\widehat{\otimes}_{\mathbb{Q}_p}\mathbb{B}^{[r,s]} = A_W\langle T/s,\ r/T\rangle,$$
$$\mathbb{A}_A^{[r,s]} = \Gamma(\mathrm{Spa}(A, A^\circ), \mathscr{A}^{[r,s]}_{\mathrm{Spa}(A,A^\circ)}) = A^\circ\widehat{\otimes}_{\mathbb{Z}_p}\mathbb{A}^{[r,s]} = A^\circ_W\langle T/s,\ r/T\rangle.$$

The following is the analogue of [Hellmann 2013, Theorem 6.9] in the nonreduced case.

**Theorem 6.4.** *Let $X$ be an adic space locally of finite type over $\mathbb{Q}_p$ and let $\mathcal{N}$ be a family of free $\varphi$-modules of rank $d$ over $\mathscr{B}_X^{[r,r_2]}$. Assume that there exists $x \in X$ and an $\mathscr{A}_X^{[r,r_2]} \otimes \kappa(x)^\circ$-lattice $N_x \subset \mathcal{N} \otimes \kappa(x)$ such that $\Phi$ induces an isomorphism*

$$\Phi : \varphi^*(N_x \otimes_{\mathscr{A}_X^{[r,r_2]}} \mathscr{A}_X^{[r,r_1]}) \xrightarrow{\sim} N_x \otimes_{\mathscr{A}_X^{[r,r_2]}} \mathscr{A}_X^{[r_1,r_2]}. \tag{6-1}$$

*Then there exists an open neighborhood $U \subset X$ of $x$ and a locally free $\mathscr{A}_U^{[r,r_2]}$-submodule $N \subset \mathcal{N}$ of rank $d$ such that*

$$N \otimes \kappa(x)^\circ = N_x,$$
$$\Phi(\varphi^* N|_{U\times\mathbb{B}_{[r,r_1]}}) = N|_{U\times\mathbb{B}_{[r_1,r_2]}},$$
$$N \otimes_{\mathscr{A}_U^{[r,r_2]}} \mathscr{B}_U^{[r,r_2]} = \mathcal{N}|_U.$$

*Proof.* We may assume that $X = \mathrm{Spa}(A, A^\circ)$ is affinoid and we may choose a Banach norm $\|\cdot\|$ and a $\mathbb{Z}_p$-subalgebra $A^+ = \{x \in A \mid \|x\| \le 1\} \subset A^\circ$ such that $A = A^+[1/p]$ and $X = \mathrm{Spa}(A, A^+) = \mathrm{Spa}(A, A^\circ)$.

Choose a basis $\underline{e}_x$ of $N_x$ and denote by $D_0 \in \mathrm{GL}_d(\mathscr{A}_X^{[r,r_2]} \otimes \kappa(x)^\circ)$ the matrix of $\Phi$ in this basis. After shrinking $X$ if necessary we may lift the matrix $D_0$ to a matrix $D$ with coefficients in $\Gamma(X, \mathscr{A}_X^{[r,r_2]})$. Localizing further we may assume that $D$ is invertible over $\Gamma(X, \mathscr{A}_X^{[r,r_2]})$, as we only need to ensure that the inverse of its determinant has coefficients $a_i \in A^+$, i.e., $\|a_i\| \le 1$ for some Banach norm $\|\cdot\|$ corresponding to $A^+$. Let us write $f \in A_w\langle T/s, r/T\rangle$ for this determinant and write $f = f^+ + f^-$ with

$$f^+ = \sum_{i\ge 0} \alpha_i \Big(\frac{T}{s}\Big)^i \in A_w\Big\langle \frac{T}{s}\Big\rangle, \qquad f^- = \sum_{i\ge 0} \beta_i \Big(\frac{r}{T}\Big)^i \in A_w\Big\langle \frac{r}{T}\Big\rangle.$$

We claim that $\alpha_i, \beta_i \in A_W^\circ$ for all $i$. But as $\alpha_i, \beta_i \xrightarrow{i\to\infty} 0$ this is clear for all but finitely many $i$. Moreover for all $i \ge 0$ we have $\alpha_i(x), \beta_i(x) \in k(x)^\circ_W$. Hence after localization on $X$ we may assume that all coefficients are integral.

Fixing a basis $\underline{b}$ of $\mathcal{N}$ we denote by $S \in \mathrm{GL}_d(\mathbb{B}_A^{[r,r]})$ the matrix of $\Phi$ in this basis. Further we denote by $V$ a lift of the change of basis matrix from the basis $\underline{e}_x$ to the basis $\underline{b} \bmod x$. From now on the proof is the same as the proof of [Hellmann 2013, Theorem 6.9]. □

**Proposition 6.5.** *Let $X = \mathrm{Spa}\,(A, A^+)$ be an affinoid adic space of finite type over $\mathbb{Q}_p$. Let $r > |\pi|$ with $r \in p^{\mathbb{Q}}$ and set $r_i = r^{1/p^i}$. Let $\mathcal{M}_r$ be a free vector bundle on $X \times \mathbb{B}_{[0,r_2]}$ together with an injection*

$$\Phi : \varphi^*(\mathcal{M}_r|_{X\times\mathbb{B}_{[0,r_1]}}) \to \mathcal{M}_r$$

*with cokernel supported at the point defined by $E(u)$. Assume that there is a free $\mathbb{A}_A^{[r,r_2]} = A^+\langle T/r_2, r/T\rangle$ submodule*

$$N_r \subset \mathcal{N}_r := \mathcal{M}_r \otimes_{\mathbb{B}_A^{[0,r_2]}} \mathbb{B}_A^{[r,r_2]}$$

*of rank $d$, containing a basis of $\mathcal{N}_r$ such that*

$$\Phi(\varphi^*(N_r \otimes_{\mathbb{A}_A^{[r,r_2]}} \mathbb{A}_A^{[r,r_1]})) = N_r \otimes_{\mathbb{A}_A^{[r,r_2]}} \mathbb{A}_A^{[r_1,r_2]}.$$

*Then fpqc-locally on $X$ there exists a free $\mathbb{A}_A^{[0,r_2]}$-submodule $M_r \subset \mathcal{M}_r$ of rank $d$, containing a basis of $\mathcal{M}_r$ such that*

$$\Phi : \varphi^*(M_r \otimes_{\mathbb{A}_A^{[0,r_2]}} \mathbb{A}_A^{[0,r_1]}) \to M_r \tag{6-2}$$

*is injective with cokernel killed by some power of $E(u)$.*

*Proof.* This is the generalization of [Hellmann 2013, Proposition 7.7] to our context. We also write $\mathcal{M}_r$ for the global sections of the vector bundle. Write $M_r' = \mathcal{M}_r \cap N_r \subset \mathcal{N}_r$. This is an $A^+\langle T/r_2\rangle$-module. Further we set

$$M_r = \mathrm{Im}(M_r' \widehat{\otimes}_{\mathbb{A}_A^{[0,r_2]}} \mathbb{A}_A^{[r,r_2]} \to \mathcal{N}_r) \cap M_r'\Big[\frac{1}{p}\Big] \subset \mathcal{N}_r.$$

Then $M_r$ is a finitely generated $A^+\langle T/r_2\rangle$-module as the ring is noetherian. First we need to make some modification in order to assure that $M_r$ is flat. Let $\mathcal{Y} = \mathrm{Spf}\, W\langle T/r_2\rangle$ denote the formal model of $\mathbb{B}_{[0,r_2]}$ and let $\mathcal{Y}' = \mathrm{Spf}\, W\langle T/r_2, r/T\rangle$ denote the formal model of $\mathbb{B}_{[r,r_2]}$. Note that $M_r[1/p] = \mathcal{M}_r$ and hence $M_r$ is rig-flat. By [Bosch and Lütkebohmert 1993, Theorem 4.1] there exists a blow-up $\widetilde{\mathcal{X}}$ of $\mathrm{Spf}(A)$ such that the strict transform $\widetilde{M}_r$ of $M_r$ in $\widetilde{\mathcal{X}} \times \mathcal{Y}$ is flat over $\widetilde{\mathcal{X}}$. We write $\mathcal{M}_{r,\widetilde{\mathcal{X}}}$ (resp. $N_{r,\widetilde{\mathcal{X}}}$) for the pullback of $\mathcal{M}_r$ (resp. $N_r$) to the generic fiber of $\widetilde{\mathcal{X}} \times \mathcal{Y}$ (resp. to $\widetilde{\mathcal{X}} \times \mathcal{Y}'$). If we set $\widetilde{M}_r' = \mathcal{M}_{r,\widetilde{\mathcal{X}}} \cap N_{r,\widetilde{\mathcal{X}}}$ then one easily finds

$$\widetilde{M}_r = (\widetilde{M}_r' \otimes_{\mathscr{A}_{\widetilde{\mathcal{X}}}^{[0,r_2]}} \mathscr{A}_{\widetilde{\mathcal{X}}}^{[r,r_2]}) \cap \widetilde{M}_r'\Big[\frac{1}{p}\Big].$$

It follows that $\widetilde{M}_r$ is stable under $\Phi$. Further, as $\widetilde{M}_r$ is flat, it has no $p$-power torsion and hence we find that the formation $(\mathcal{M}_{r,\widetilde{\mathcal{X}}}, N_{r,\widetilde{\mathcal{X}}}) \mapsto \widetilde{M}_r$ commutes with base change $\mathrm{Spf}\,\mathcal{O} \hookrightarrow \widetilde{\mathcal{X}}$ for any finite flat $\mathbb{Z}_p$-algebra $\mathcal{O}$; compare the proof of [Hellmann 2013, Proposition 7.7]. Especially this pullback is free over $\mathcal{O} \otimes_{\mathbb{Z}_p} W\langle T/r_2\rangle$ and the cokernel of $\Phi$ is annihilated by $E(u)^{k_\mu}$ for some $k_\mu \gg 0$ depending only on the Hodge polygon $\mu$ (for an arbitrary finite flat $\mathbb{Z}_p$-algebra $\mathcal{O}$ this follows by forgetting the $\mathcal{O}$-structure and only considering the $\mathbb{Z}_p$-structure).

It follows that the restriction of $\widetilde{M}_r$ to the reduced special fiber $\widetilde{\mathcal{X}}_0$ of $\widetilde{\mathcal{X}}$ is locally free over $\widetilde{\mathcal{X}}_0 \times \mathbb{A}^1$ and hence, as in the proof of [Hellmann 2013, Proposition 7.7] we may locally lift a basis and find that $\widetilde{M}_r$ is locally on $\widetilde{\mathcal{X}}$ free over $\widetilde{\mathcal{X}} \times \mathcal{Y}$.

It is only left to show that $E(u)^{k_\mu}\operatorname{coker}\Phi = 0$ over $\widetilde{\mathcal{X}}$. To do so we may localize and assume that $\widetilde{\mathcal{X}}$ is affine. By abuse of language we denote it again by $\operatorname{Spf} A^+$ and write $N = E(u)^{k_\mu}\operatorname{coker}\Phi$. If $I$ denotes the ideal of nilpotent elements in $A^+$, we need to show that the multiplication $I \otimes_{A^+} N \to N$ is the zero map. Indeed, if $IN = 0$, then $N$ does not change if we pull back the situation to the reduced ring $A^+/I$. However, for $A/I$ Nakayama's lemma implies that $E(u)^{k_\mu}\operatorname{coker}\Phi$ vanishes if it vanishes after all possible pullbacks $\operatorname{Spf}\mathcal{O} \hookrightarrow \tilde{\mathcal{X}}$. We already remarked above that in the case $A^+ = \mathcal{O}$ a finite flat $\mathbb{Z}_p$-algebra the cokernel is killed by $E(u)^{k_\mu}$.

We now show that $IN = 0$. For some $k \gg 0$ we know that $I^k \otimes_{A^+} N \to N$ is the zero map, as $I$ is nilpotent. Then $N = N/I^k$ and the multiplication map $I^{k-1}\otimes_{A^+} N \to N$ factors over $I^{k-1}/I^k \otimes_{A^+} N \to N$ and this is a map of finitely generated $A^+/I^k$-modules which vanishes after pulling back to a quotient $A^+/I^k \to \mathcal{O}_L$ onto the ring of integers in some finite extension $L$ of $\mathbb{Q}_p$. This can be seen as follows. The map on this pullback is induced by the pullback of the multiplication to a quotient of $A^+$ which is finite flat over $\mathbb{Z}_p$, and where $N$ is known to vanish by the above. It follows that $I^{k-1} \otimes_{A^+} N \to N$ is the zero map and by descending induction we find that $I$ acts trivially on $N$. □

*Proof of Theorem 6.3.* Fix some $r > |\pi|$ and redefine

$$X^{\mathrm{int}} = \left\{ x \in X \;\middle|\; \begin{array}{c} \mathcal{M}|_{X\times\mathbb{B}_{[r,r_2]}} \otimes \kappa(x) \text{ contains an } \mathscr{A}_X^{[r,r_2]} \otimes \kappa(x)^\circ \text{ lattice } N_x \\ \text{such that } \Phi \text{ induces an isomorphism} \\ \varphi^*(N_x \otimes_{\mathscr{A}_X^{[r,r_2]}} \mathscr{A}_X^{[r,r_1]}) \xrightarrow{\sim} N_x \otimes_{\mathscr{A}_X^{[r,r_2]}} \mathscr{A}_X^{[r_1,r_2]} \end{array} \right\}.$$

By Theorem 6.4 this subset is open and we need to show that the restriction $\mathcal{M}|_{X^{\mathrm{int}}}$ is étale. Then it follows directly that $X^{\mathrm{int}}$ coincides with the characterization in the theorem, as the notion of being étale at points may be checked fpqc-locally by [Hellmann 2013, Proposition 6.14].

However Proposition 6.5 provides (locally on $X^{\mathrm{int}}$) an integral model $\mathfrak{M}_{[0,r_2]}$ over $X \times \mathbb{B}_{[0,r_2]}$. Now we can glue $\mathfrak{M}_{[0,r_2]}$ and $\varphi^*\mathfrak{M}_{[0,r_2]}$ over $X \times \mathbb{B}_{[r_2,r_3]}$ along the isomorphism $\Phi$. Hence we can extend $\mathfrak{M}_{[0,r_2]}$ to a model $\mathfrak{M}_{[0,r_3]}$ over $X \times \mathbb{B}_{[0,r_3]}$. Proceeding by induction we get a model $\mathfrak{M}$ on $X \times \mathbb{U}$ and [Hellmann 2013, Proposition 6.5] guarantees that $\mathfrak{M}$ is locally in $X$ free over $\mathscr{A}_X^{[0,1)}$ (it is assumed $\mathcal{N}$ is free in [loc. cit.]. However, its proof only uses the fact that the restriction of $\mathcal{N}$ to an annulus $X \times \mathbb{B}_{[r,r^{1/p^2}]}$ is free. This is always true after localizing on $X$; see [Lütkebohmert 1977]). Hence it is the desired étale model. □

**Corollary 6.6.** *Let $\mu$ be a cocharacter as in* (2-5) *with reflex field $E_\mu$. Then there is an open substack $\mathscr{H}^{\mathrm{ad,int}}_{\varphi,N,\preceq\mu} \subset \mathscr{H}^{\mathrm{ad}}_{\varphi,N,\preceq\mu}$ such that $f : X \to \mathscr{H}^{\mathrm{ad}}_{\varphi,N,\preceq\mu}$ factors over $\mathscr{H}^{\mathrm{ad,int}}_{\varphi,N,\preceq\mu}$ if and only if the family $(\mathcal{M}, \Phi_{\mathcal{M}}, N_\nabla^{\mathcal{M}})$ defined by $f$ and $\underline{\mathcal{M}}$ is étale.*

*Proof.* Let $\underline{\mathcal{M}}(\underline{D})$ be the universal $(\varphi, N_\nabla)$-module over $\mathscr{H}^{\mathrm{ad}}_{\varphi,N,\preceq\mu}$. By Theorem 6.3 the set $\mathscr{H}^{\mathrm{ad,int}}_{\varphi,N,\preceq\mu} := \{x \in \mathscr{H}^{\mathrm{ad}}_{\varphi,N,\preceq\mu} : \underline{\mathcal{M}}(\underline{D}) \text{ is étale at } x\}$ is open and above it $\underline{\mathcal{M}}(\underline{D})$ is étale. If $f$ factors over $\mathscr{H}^{\mathrm{ad,int}}_{\varphi,N,\preceq\mu}$ then $(\mathcal{M}, \Phi_{\mathcal{M}}, N_\nabla^{\mathcal{M}})$ is the pullback of the universal $\underline{\mathcal{M}}(\underline{D})$ and hence is étale. Conversely if $(\mathcal{M}, \Phi_{\mathcal{M}}, N_\nabla^{\mathcal{M}})$ is étale, then it is étale at all points and $f$ factors over $\mathscr{H}^{\mathrm{ad,int}}_{\varphi,N,\preceq\mu}$, because the notion of being étale at points may be checked fpqc-locally by [Hellmann 2013, Proposition 6.14]. □

**Proposition 6.7.** *Let $L$ be a finite extension of $E_\mu$, then $\mathscr{H}^{\mathrm{ad,int}}_{\varphi,N,\preceq\mu}(L) = \mathscr{H}^{\mathrm{ad,wa}}_{\varphi,N,\preceq\mu}(L)$ and hence $\mathscr{H}^{\mathrm{ad,int}}_{\varphi,N,\preceq\mu} \subset \mathscr{H}^{\mathrm{ad,wa}}_{\varphi,N,\preceq\mu}$.*

*Proof.* We show that being weakly admissible translates into being pure of slope zero over the Robba ring (in the sense of [Kedlaya 2008]) under the equivalence of categories from Theorem 4.6. However, the proof is the same as in [Kisin 2006, Theorem 1.3.8]. One easily verifies that the functor $\underline{\mathcal{M}}$ preserves the slope and that the slope filtration on the base change of $\underline{\mathcal{M}}(D, \Phi, N, \mathfrak{q})$ to the Robba ring extends to all of $\underline{\mathcal{M}}(D, \Phi, N, \mathfrak{q})$. Compare [Kisin 2006, Proposition 1.3.7].

As in [Hellmann 2013, Theorem 7.6 (ii)] the second part is now a consequence of the fact that $\mathscr{H}^{\mathrm{ad,wa}}_{\varphi,N,\preceq\mu} \subset \mathscr{H}^{\mathrm{ad}}_{\varphi,N,\preceq\mu}$ is the maximal open subspace whose rigid analytic points are exactly the weakly admissible ones, see Lemma 5.10. □

Pappas and Rapoport [2009, 5.b] defined a *period morphism* from a stack of integral data to a stack of filtered $\varphi$-modules as follows. Let $d > 0$ and let $\mu : \mathbb{G}_{m,\bar{\mathbb{Q}}_p} \to \widetilde{T}_{\bar{\mathbb{Q}}_p}$ be a cocharacter as in (2-5). Pappas and Rapoport [2009, 3.d] defined an fpqc-stack $\widehat{\mathcal{C}}_{\mu,K}$ on the category $\mathrm{Nil}_{\mathcal{O}_{E_\mu}}$ of schemes over the ring of integers $\mathcal{O}_{E_\mu}$ of $E_\mu$ on which $p$ is locally nilpotent. If $R$ is an $\mathcal{O}_{E_\mu}$-algebra, we set $R_W = R \otimes_{\mathbb{Z}_p} W$ and denote by $\varphi : R_W(\!(u)\!) \to R_W(\!(u)\!)$ the ring homomorphism that is the identity on $R$, the $p$-Frobenius on $W$ and that maps $u$ to $u^p$. Now the $R$-valued points of the stack $\widehat{\mathcal{C}}_{\mu,K}$ are given by a subset

$$\widehat{\mathcal{C}}_{\mu,K}(R) \subset \{\mathfrak{M},\ \Phi : \varphi^*\mathfrak{M}[1/u] \xrightarrow{\sim} \mathfrak{M}[1/u]\},$$

where $\mathfrak{M}$ is an $R_W[\![u]\!] = (R \otimes_{\mathbb{Z}_p} W)[\![u]\!]$-module that is fpqc-locally on $\operatorname{Spec} R$ free as an $R_W[\![u]\!]$-module of rank $d$. This subset is cut out by a condition prescribing the relative position of $\Phi(\varphi^*\mathfrak{M})$ with respect to $\mathfrak{M}$ at the locus $E(u) = 0$ in terms of the cocharacter $\mu$; see [Pappas and Rapoport 2009, 3.c,d] for the precise definition.

If $\mu$ is minuscule they defined a *period map*

$$\Pi(\mathcal{X}) : \widehat{\mathcal{C}}_{\mu,K}(\mathcal{X}) \to \mathscr{D}^{\mathrm{ad}}_{\varphi,\mu}(\mathcal{X}^{\mathrm{rig}});$$

see [Pappas and Rapoport 2009, (5.37)]. Note that $\widehat{\mathcal{C}}_{\mu,K}$ is a substack of $\widehat{\mathcal{C}}_{d,K}$ of [loc. cit.] if and only if $\mu$ is minuscule. Moreover, the period morphism of [loc. cit.] maps the closed substack $\widehat{\mathcal{C}}_{\mu,K}$ to the corresponding closed substack $\mathscr{D}^{\mathrm{ad}}_{\varphi,\mu}$ of their target $\mathscr{D}_{d,K}$.

If $\mu$ is not miniscule we cannot hope for a period map with target $\mathscr{D}^{\mathrm{ad}}_{\varphi,\mu}$. However, if we replace the target by $\mathscr{H}^{\mathrm{ad}}_{\varphi,\preceq\mu}$, then we can again define a period map as follows (note that $\mathscr{D}^{\mathrm{ad}}_{\varphi,\mu} = \mathscr{H}^{\mathrm{ad}}_{\varphi,\preceq\mu}$ if $\mu$ is miniscule). Let $R$ be a $p$-adically complete $\mathcal{O}_{E_\mu}$-algebra topologically of finite type over $\mathcal{O}_{E_\mu}$ and let $(\mathfrak{M}, \Phi) \in \widehat{\mathcal{C}}_{\mu,K}(\operatorname{Spf} R)$. The construction of Section 4 associates to

$$(\mathcal{M}, \Phi_{\mathcal{M}}) = (\mathfrak{M}, \Phi) \otimes_{R_W[\![u]\!]} \mathscr{B}^{[0,1)}_{\mathrm{Spa}(R[1/p],R)}, \tag{6-3}$$

a $\varphi$-module with Hodge–Pink lattice over $\mathrm{Spa}(R[1/p], R)$. Given a formal scheme $\mathcal{X}$ locally topologically of finite type over $\mathcal{O}_{E_\mu}$, this yields a period functor

$$\Pi(\mathcal{X}) : \widehat{\mathcal{C}}_{\mu,K}(\mathcal{X}) \to \mathscr{H}^{\mathrm{ad}}_{\varphi,\preceq\mu}(\mathcal{X}^{\mathrm{rig}}), \tag{6-4}$$

where $\mathcal{X}^{\mathrm{rig}}$ denotes the generic fiber of the formal scheme $\mathcal{X}$ in the sense of rigid geometry (or in the sense of adic spaces). We point out that we cannot define a period map mapping to $\mathscr{D}^{\mathrm{ad}}_{\varphi,\mu}$ if $\mu$ is not miniscule, as the family of vector bundles on the open unit disc defined by (6-3) is not necessarily associated to a filtered $\varphi$-module: the monodromy operator $N_\nabla^{\mathcal{M}}$ is not necessarily holomorphic. When $\mathcal{X} = \operatorname{Spf}\mathcal{O}_L$ for a finite field extension $L$ of $E_\mu$, it was shown by Genestier and Lafforgue [2012, Théorème 0.6] that $\Pi(\operatorname{Spf}\mathcal{O}_L)\otimes_{\mathbb{Z}_p}\mathbb{Q}_p$ is fully faithful, and surjective onto $\mathscr{H}^{\mathrm{ad,wa}}_{\varphi,\preceq\mu}(L) = \mathscr{H}^{\mathrm{ad,int}}_{\varphi,\preceq\mu}(L)$.

**Remark 6.8.** From the point of view of Galois representations it is not surprising that we cannot define a general period morphism using filtered $\varphi$-modules. If $R$ is finite over $\mathcal{O}_{E_\mu}$, then the points of $\widehat{\mathcal{C}}_{\mu,K}(R)$ correspond to $\mathscr{G}_{K_\infty}$-representations rather than to $\mathscr{G}_K$-representations. This also explains why the target of the period map is $\mathscr{H}^{\mathrm{ad}}_{\varphi,\preceq\mu}$ instead of $\mathscr{H}^{\mathrm{ad}}_{\varphi,N,\preceq\mu}$: the $\mathscr{G}_{K_\infty}$-representation does not see the monodromy.

If we want to take the monodromy into account we have to consider a stack $\widehat{\mathcal{C}}_{\mu,N,K}$ whose $\mathcal{X}$-valued points are given by $(\mathfrak{M}, \Phi, N)$ with $(\mathfrak{M}, \Phi) \in \widehat{\mathcal{C}}_{\mu,K}(\mathcal{X})$ and $N : \mathfrak{M}/u\mathfrak{M} \to \mathfrak{M}/u\mathfrak{M}$ satisfying

$$N \circ \overline{\Phi}(n) = p \cdot \overline{\Phi}(n) \circ N. \tag{6-5}$$

Here $(\mathfrak{M}(n), \Phi(n)) = (\mathfrak{M}, \Phi) \otimes_{W[\![u]\!]} \mathbb{A}^{[0,1)}(n)$ is the twist of $(\mathfrak{M}, \Phi)$ with the object $\mathbb{A}^{[0,1)}(n)$ defined before Definition 6.2 and $n \gg 0$ is some integer such that $\Phi(n)(\varphi^*\mathfrak{M}) \subset \mathfrak{M}$ and $\overline{\Phi}$ denotes the reduction of $\Phi$ modulo $u$. Note that given $\mu$ we may choose an $n$ like that for all $(\mathfrak{M}, \Phi) \in \widehat{\mathcal{C}}_{\mu,K}(\mathcal{X})$ and the map $\overline{\Phi}$ (and hence the equation (6-5)) makes sense after this twist. Further the condition defined by (6-5) is independent of the chosen $n$.

**Remark 6.9.** (i) Using (2-7) we observe that if $\mu_{\psi,d} \geq 0$ for all $\psi$, and if $L$ is a finite extension of $E_\mu$, a $\operatorname{Spf}\mathcal{O}_L$-valued point of the stack $\widehat{\mathcal{C}}_{\mu,N,K}$ gives rise to an object of the category $\mathrm{Mod}^{\varphi,N}_{/\mathfrak{S}}$ in the sense of [Kisin 2006, (1.3.12)]. We only use the twist in order to define the stack in the general case (i.e., if $\Phi(\varphi^*\mathfrak{M})$ is not contained in $\mathfrak{M}$). Kisin's definition takes place in the generic fiber. However, we can not use this as a good definition as our stack is defined for $p$-power torsion objects.

(ii) Note that we do not know much about the stack $\widehat{\mathcal{C}}_{\mu,N,K}$ and its definition is rather ad hoc. Especially we doubt that it is flat over $\operatorname{Spf}\mathbb{Z}_p$. This means that there is no reason to expect that we can reconstruct Kisin's semistable deformation rings [2008] by using a similar construction to that in [Pappas and Rapoport 2009, §4].

In this general case described above we obtain a similar period morphism

$$\widehat{\mathcal{C}}_{\mu,N,K}(\mathcal{X}) \to \mathscr{H}^{\mathrm{ad}}_{\varphi,N,\preceq\mu}(\mathcal{X}^{\mathrm{rig}}). \tag{6-6}$$

As in [Hellmann 2013, Theorem 7.8] the above allows us to determine the image of the period morphism. Recall that a valued field $(L, v_L)$ over $\mathbb{Q}_p$ is called *of $p$-adic type* if it is complete, topologically finitely generated over $\mathbb{Q}_p$ and if for all $f_1, \ldots, f_m \in L$ the closure of $\mathbb{Q}_p[f_1 \ldots, f_m]$ inside $L$ is a Tate algebra, i.e., the quotient of some $\mathbb{Q}_p\langle T_1, \ldots, T_{m'}\rangle$.

**Corollary 6.10.** *The substack $\mathscr{H}^{\mathrm{ad,int}}_{\varphi,N,\preceq\mu}$ is the image of the period morphism* (6-6) *in the following sense*:

(i) *If $\mathcal{X}$ is a $p$-adic formal scheme and $(\mathfrak{M},\Phi,N)\in\widehat{\mathcal{C}}_{\mu,N,K}(X)$, then $\Pi(\mathcal{X})(\mathfrak{M},\Phi,N)\in\mathscr{H}^{\mathrm{ad,int}}_{\varphi,N,\preceq\mu}(\mathcal{X}^{\mathrm{rig}})$.*

(ii) *Let $L$ be a field of $p$-adic type over $E_\mu$ and $(D,\Phi,N,\mathfrak{q})\in\mathscr{H}_{\varphi,N,\preceq\mu}(L)$. Then there exists $(\mathfrak{M},\Phi,N)\in\widehat{\mathcal{C}}_{\mu,N,K}(\mathrm{Spf}\ L^+)$ such that $\Pi(\mathrm{Spf}\ L^+)(\mathfrak{M},\Phi,N)=(D,\Phi,N,\mathfrak{q})$ if and only if*

$$\underline{\mathcal{M}}(D)=\mathfrak{M}\otimes_{L^+_W[\![u]\!]}\mathscr{B}_L^{[0,1)}.$$

*is étale, if and only if $\mathrm{Spa}(L,L^+)\to\mathscr{H}^{\mathrm{ad}}_{\varphi,N,\preceq\mu}$ factors over $\mathscr{H}^{\mathrm{ad,int}}_{\varphi,N,\preceq\mu}$.*

(iii) *Let $X\in\mathrm{Ad}^{\mathrm{lft}}_{E_\mu}$ and let $f:X\to\mathscr{H}^{\mathrm{ad}}_{\varphi,N,\preceq\mu}$ be a morphism defined by $(D,\Phi,N,\mathfrak{q})$. Then $f$ factors over $\mathscr{H}^{\mathrm{ad,int}}_{\varphi,N,\preceq\mu}$ if and only if there exists a fpqc-covering $(U_i\to X)_{i\in I}$ and formal models $\mathcal{U}_i$ of $U_i$ together with $(\mathfrak{M}_i,\Phi_i,N)\in\widehat{\mathcal{C}}_{\mu,N,K}(\mathcal{U}_i)$ such that $\Pi(\mathcal{U}_i)(\mathfrak{M}_i,\Phi_i,N)=(D,\Phi,N,\mathfrak{q})|_{U_i}$.*

**Remark 6.11.** If we consider the period morphism without monodromy, then we obtain a similar characterization of the stack $\mathscr{H}^{\mathrm{ad,int}}_{\varphi,\preceq\mu}\subset\mathscr{H}^{\mathrm{ad}}_{\varphi,\preceq\mu}$ as the image of the period morphism (6-4).

## 7. Sheaves of period rings and the admissible locus

We recall the definition of some sheafified period rings from [Hellmann 2013]. In doing so we will also correct mistakes in [loc. cit.] (in particular the proofs of Corollary 8.8, the definition of a family of crystalline representations, and the proof of Proposition 8.24 in [Hellmann 2013]).

Let $R=\varprojlim\mathcal{O}_{\mathbb{C}_p}/p\mathcal{O}_{\mathbb{C}_p}$ be the inverse limit with transition maps given by the $p$-th power. Given a reduced $p$-adically complete $\mathbb{Z}_p$-algebra $A^+$ topologically of finite type, we define

$$A^+\widehat{\otimes}_{\mathbb{Z}_p}W(R)=\varprojlim_i A^+\widehat{\otimes}_{\mathbb{Z}_p}W_i(R),$$

where the completed tensor product on the right-hand side means completion with respect to the canonical topology on the truncated Witt vectors $W_i(R)$ and the discrete topology on $A^+/p^iA^+$.

If $X$ is a reduced adic space locally of finite type over $\mathbb{Q}_p$, then there are sheaves $\mathcal{O}_X^+\widehat{\otimes}W(R)$ and $\mathcal{O}_X\widehat{\otimes}W(R)$ whose sections over an affinoid open $U=\mathrm{Spa}(A,A^+)\subset X$ are given by

$$\Gamma(U,\mathcal{O}_X^+\widehat{\otimes}W(R))=A^+\widehat{\otimes}_{\mathbb{Z}_p}W(R),$$
$$\Gamma(U,\mathcal{O}_X\widehat{\otimes}W(R))=\big(A^+\widehat{\otimes}_{\mathbb{Z}_p}W(R)\big)\Big[\frac{1}{p}\Big].$$

In the same fashion we can define sheaves of topological rings $\mathcal{O}_X^+\widehat{\otimes}W(\mathrm{Frac}\,R)$ and $\mathcal{O}_X\widehat{\otimes}W(\mathrm{Frac}\,R)$.

Let $\mathbb{A}^{[0,1)}=W[\![u]\!]$ and let $\mathbb{A}$ denote the $p$-adic completion of $W(\!(u)\!)$. Further let $\mathbb{B}=\mathbb{A}[1/p]$. We fix an element $\pi^\flat=(\pi_n)_n\in R$ with $\pi_0=\pi$. Depending on this element there are embeddings of $\mathbb{A}^{[0,1)}$, $\mathbb{A}$ and $\mathbb{B}$ into $W(\mathrm{Frac}\,R)[1/p]$ sending $u$ to the Teichmüller representative $[\pi^\flat]\in W(R)$ of $\pi^\flat$. We write $\widetilde{\mathbb{A}}$ for the ring of integers in the completion $\widetilde{\mathbb{B}}$ of the maximal unramified extension of $\mathbb{B}$ inside $W(\mathrm{Frac}\,R)[1/p]$. Finally we set $\widetilde{\mathbb{A}}^{[0,1)}=\widetilde{\mathbb{A}}\cap W(R)\subset W(\mathrm{Frac}\,R)$. All these rings come along with a Frobenius endomorphism $\varphi$ which is induced by the canonical Frobenius on $W(\mathrm{Frac}\,R)$. Note that all these rings have a canonical topology induced from the one on $W(\mathrm{Frac}\,R)$.

**Remark 7.1.** We warn (and apologize to) the reader that the notations used in this paragraph do often not agree with the notations that are nowadays standard in $p$-adic Hodge theory. However, we often refer to [Hellmann 2013] and it seems to cause less confusion using the notations used there.

We define sheafified versions of these rings as follows; compare [Hellmann 2013, 8.1]. Let $X$ be a reduced adic space locally of finite type over $\mathbb{Q}_p$. We define the sheaves $\mathscr{A}_X$, $\widetilde{\mathscr{A}}_X$, $\mathscr{A}_X^{[0,1)}$ and $\widetilde{\mathscr{A}}_X^{[0,1)}$ by specifying their sections on open affinoids $U = \mathrm{Spa}(A, A^+) \subset X$: we define $\Gamma(U, \mathscr{A}_X)$, $\Gamma(U, \widetilde{\mathscr{A}}_X)$, $\Gamma(U, \mathscr{A}_X^{[0,1)})$ and $\Gamma(U, \widetilde{\mathscr{A}}_X^{[0,1)})$ to be the closures (with respect to the natural, i.e., $(p, [\pi^\flat])$-adic, topology) of $A^+ \otimes_{\mathbb{Z}_p} \mathbb{A}$, $A^+ \otimes_{\mathbb{Z}_p} \widetilde{\mathbb{A}}$, $A^+ \otimes_{\mathbb{Z}_p} \mathbb{A}^{[0,1)}$ and $A^+ \otimes_{\mathbb{Z}_p} \widetilde{\mathbb{A}}^{[0,1)}$ in $\Gamma(U, \mathcal{O}_X^+ \widehat{\otimes} W(\mathrm{Frac}\, R))$, respectively.

Further we consider the rational analogues $\mathscr{B}_X$, $\widetilde{\mathscr{B}}_X$, $\mathscr{B}_X^{[0,1)}$ and $\widetilde{\mathscr{B}}_X^{[0,1)}$ of these sheaves given by inverting $p$ in $\mathscr{A}_X$, $\widetilde{\mathscr{A}}_X$, $\mathscr{A}_X^{[0,1)}$ and $\widetilde{\mathscr{A}}_X^{[0,1)}$, respectively.

Finally we recall the construction of the sheaf $\mathcal{O}_X \widehat{\otimes} B_{\mathrm{cris}}$ from [loc. cit., 8.1]. For a reduced adic space $X$ the map $\theta : W(R) \to \mathcal{O}_{\mathbb{C}_p}$ given by $[(x, x^{1/p}, x^{1/p^2}, \dots)] \mapsto x$ extends to an $\mathcal{O}_X^+$-linear map

$$\theta_X : \mathcal{O}_X^+ \widehat{\otimes} W(R) \to \mathcal{O}_X^+ \widehat{\otimes} \mathcal{O}_{\mathbb{C}_p},$$

where the completed tensor product denotes the $p$-adic completion. We define $\mathcal{O}_X^+ \widehat{\otimes} A_{\mathrm{cris}}$ to be the $p$-adic completion of the divided power envelope of $\mathcal{O}_X^+ \widehat{\otimes} W(R)$ with respect to the kernel of $\theta_X$. We claim that $\mathcal{O}_X^+ \widehat{\otimes} A_{\mathrm{cris}}$ equals the $p$-adic completion of the tensor product $\mathcal{O}_X^+ \otimes_{\mathbb{Z}_p} A_{\mathrm{cris}}$. Namely, the kernel of $\theta_X$ is generated by the kernel of $\theta$. The latter in turn is generated by $p - [p^\flat]$, where $p^\flat = (x_n)_n \in R$ with $x_0 = p$ and $[\,\cdot\,]$ denotes the Teichmüller lift. Therefore, the divided power envelope is constructed by adjoining $(p - [p^\flat])^n / n!$ for all $n \in \mathbb{N}$, and this proves our claim. Finally we set

$$\begin{aligned}
\mathcal{O}_X \widehat{\otimes} B_{\mathrm{cris}}^+ &= \big(\mathcal{O}_X^+ \widehat{\otimes} A_{\mathrm{cris}}\big)[1/p], \\
\mathcal{O}_X \widehat{\otimes} B_{\mathrm{cris}} &= \big(\mathcal{O}_X \widehat{\otimes} B_{\mathrm{cris}}^+\big)[1/t], \\
\mathcal{O}_X \widehat{\otimes} B_{\mathrm{st}}^+ &= \big(\mathcal{O}_X \widehat{\otimes} B_{\mathrm{cris}}^+\big)[\ell_u], \\
\mathcal{O}_X \widehat{\otimes} B_{\mathrm{st}} &= \big(\mathcal{O}_X \widehat{\otimes} B_{\mathrm{cris}}\big)[\ell_u].
\end{aligned}$$

Here $t = \log[(1, \varepsilon_1, \varepsilon_2, \dots)] \in B_{\mathrm{cris}}$ is the period of the cyclotomic character (where $(\varepsilon_i)$ is a compatible system of $p^i$-th roots of unity) and $\ell_u$ is an indeterminate thought of as a formal logarithm of $[\pi^\flat]$.

**Remark 7.2.** The indeterminate $\ell_u$ considered here is the same indeterminate as in section 2.2.(b) and we identify both indeterminates. That is, the inclusion $\mathbb{B}^{[0,1)} \subset B_{\mathrm{cris}}^+$ given by $u \mapsto [\pi^\flat]$ will be extended to $\mathbb{B}^{[0,1)}[\ell_u] \hookrightarrow B_{\mathrm{st}}^+$ by means of $\ell_u \mapsto \ell_u$ and similarly for the sheafified versions.

**Lemma 7.3.** *Let $Y = \mathrm{Spa}(B, B^+)$ be an reduced adic space that is finite over $X = \mathrm{Spa}(A, A^+)$. Then we have canonical isomorphisms*

$$\begin{aligned}
\widetilde{\mathscr{B}}_Y &\cong \widetilde{\mathscr{B}}_X \otimes_{\mathcal{O}_X} \mathcal{O}_Y, & \mathscr{B}_Y^{[0,1)} &\cong \mathscr{B}_X^{[0,1)} \otimes_{\mathcal{O}_X} \mathcal{O}_Y, \\
\mathcal{O}_Y \widehat{\otimes} B_{\mathrm{cris}}^+ &\cong (\mathcal{O}_X \widehat{\otimes} B_{\mathrm{cris}}^+) \otimes_{\mathcal{O}_X} \mathcal{O}_Y, & \mathcal{O}_Y \widehat{\otimes} B_{\mathrm{cris}} &\cong (\mathcal{O}_X \widehat{\otimes} B_{\mathrm{cris}}) \otimes_{\mathcal{O}_X} \mathcal{O}_Y, \\
\mathcal{O}_Y \widehat{\otimes} B_{\mathrm{st}}^+ &\cong (\mathcal{O}_X \widehat{\otimes} B_{\mathrm{st}}^+) \otimes_{\mathcal{O}_X} \mathcal{O}_Y, & \mathcal{O}_Y \widehat{\otimes} B_{\mathrm{st}} &\cong (\mathcal{O}_X \widehat{\otimes} B_{\mathrm{st}}) \otimes_{\mathcal{O}_X} \mathcal{O}_Y.
\end{aligned}$$

*Proof.* This is a direct consequence of the construction (and the fact that we do not have to complete tensor products with finitely generated modules). □

We can consider these sheaves also on nonreduced spaces by locally embedding $X$ into a reduced space $Y$ and restricting the corresponding sheaves from $Y$ to $X$, i.e., by applying $-\otimes_{\mathcal{O}_Y}\mathcal{O}_X$. Thanks to the above lemma, the sheaves like $\mathcal{O}_X\widehat{\otimes}B^+_{\mathrm{cris}}$ then do not depend on the choice of an embedding. With this definition the claim of Lemma 7.3 also holds true for nonreduced adic spaces.

**Remark 7.4.** For nonreduced spaces we make this slightly involved definition for the following reason: the construction of rings like $A_{\mathrm{cris}}$ involves a $p$-adic completion. But the rings of integral elements $A^+$ for an adic space $\mathrm{Spa}(A, A^+)$ (i.e., the power bounded elements in $A$) are not $p$-adically complete: their $p$-adic completion would kill the nilpotent elements!

On $\mathcal{O}_X\widehat{\otimes}B_{\mathrm{cris}}$ there is a canonical Frobenius $\varphi$ induced by the Frobenius on $\mathcal{O}^+_X\widehat{\otimes}W(R)$. This endomorphism extends to a morphism

$$\varphi : \mathcal{O}_X\widehat{\otimes}B_{\mathrm{st}} \to \mathcal{O}_X\widehat{\otimes}B_{\mathrm{st}},$$

where $\varphi(\ell_u) = p\ell_u$. Further $N = d/d\ell_u$ defines an endomorphism of $\mathcal{O}_X\widehat{\otimes}B_{\mathrm{st}}$ which satisfies $N\varphi = p\varphi N$.

Finally the continuous $\mathcal{G}_K$-action on $\mathcal{O}^+_X\widehat{\otimes}W(R)$ extends to $\mathcal{O}_X\widehat{\otimes}B_{\mathrm{cris}}$ and we further extend this action to $\mathcal{O}_X\widehat{\otimes}B_{\mathrm{st}}$ by means of $\gamma\cdot\ell_u = \ell_u + c(\gamma)t$, where $c : \mathcal{G}_K \to \mathbb{Z}_p$ is defined by $\gamma(\pi_n) = \pi_n\cdot(\varepsilon_n)^{c(\gamma)}$ for all $n \geq 0$.

**Lemma 7.5.** *Let $Y = \mathrm{Spa}(A, A^+)$ be an adic space locally of finite type over $\mathbb{Q}_p$.*

(a) *Let $g \in \Gamma(Y, \mathcal{O}_Y\widehat{\otimes}B^+_{\mathrm{cris}})$. Then $g \in \Gamma(Y, \mathcal{O}_Y) \subset \Gamma(Y, \mathcal{O}_Y\widehat{\otimes}B^+_{\mathrm{cris}})$ if and only if for every quotient $A \twoheadrightarrow A'$ onto a finite-dimensional $\mathbb{Q}_p$-algebra $A'$ the element*

$$g\otimes 1 \in \Gamma(Y, \mathcal{O}_Y\widehat{\otimes}B^+_{\mathrm{cris}})\otimes_A A' \cong A'\otimes B^+_{\mathrm{cris}}$$

*actually lies in $A' \subset A'\otimes_{\mathbb{Q}_p} B^+_{\mathrm{cris}}$.*

(b) *Let $g \in \Gamma(Y, \widetilde{\mathscr{B}}_Y)$. Then $g \in \Gamma(Y, \mathcal{O}_Y) \subset \Gamma(Y, \widetilde{\mathscr{B}}_Y)$ if and only if for every quotient $A \twoheadrightarrow A'$ onto a finite-dimensional $\mathbb{Q}_p$-algebra $A'$ the element*

$$g\otimes 1 \in \Gamma(\mathrm{Spa}(A', A'^+), \widetilde{\mathscr{B}}_Y\otimes_A A') \cong A'\otimes_{\mathbb{Q}_p}\widetilde{\mathbb{B}}$$

*actually lies in $A' \subset A'\otimes_{\mathbb{Q}_p}\widetilde{\mathbb{B}}$.*

(c) *Assume that $A$ is reduced. Let $g \in \Gamma(Y, \widetilde{\mathscr{A}}^{[0,1)}_Y)$. Then $g \in \Gamma(Y, \mathscr{A}^{[0,1)}_Y) \subset \Gamma(Y, \widetilde{\mathscr{A}}^{[0,1)}_Y)$ if and only if for every rigid analytic point $y \in Y$ the element $g(y) := g\otimes_{A^+}\kappa(y)^+ \in \kappa(y)^+\widehat{\otimes}_{\mathbb{Z}_p}\widetilde{\mathbb{A}}^{[0,1)}$ actually lies in $\kappa(y)^+\widehat{\otimes}_{\mathbb{Z}_p}\mathbb{A}^{[0,1)} \subset \kappa(y)^+\widehat{\otimes}_{\mathbb{Z}_p}\widetilde{\mathbb{A}}^{[0,1)}$.*

Note that the identifications

$$\Gamma(\mathrm{Spa}(A', A'^+), (\widetilde{\mathscr{B}}_Y)\otimes_A A') \cong A'\otimes_{\mathbb{Q}_p}\widetilde{\mathbb{B}} \quad\text{and}\quad \Gamma(Y, \mathcal{O}_Y\widehat{\otimes}B^+_{\mathrm{cris}})\otimes_A A' \cong A'\otimes B^+_{\mathrm{cris}}$$

used in the formulation of the lemma are a direct consequence of Lemma 7.3 and the remark following it.

*Proof.* (a) Clearly the condition is necessary. We now show that it is sufficient. Let us choose a closed immersion $Y = \mathrm{Spa}(A, A^+) \hookrightarrow X = \mathrm{Spa}(C, C^+)$ with $C$ a reduced Tate ring topologically of finite type over $\mathbb{Q}_p$. Then $C \twoheadrightarrow A$ and our definitions imply that $A\widehat{\otimes}B^+_{\mathrm{cris}}$ is the quotient of $C\widehat{\otimes}B^+_{\mathrm{cris}}$ by the kernel of $C \to A$. We choose elements $b_i \in A_{\mathrm{cris}}$ with $b_0 = 1$ whose images $\bar{b}_i$ in $A_{\mathrm{cris}}/pA_{\mathrm{cris}}$ form an $\mathbb{F}_p$-basis of $A_{\mathrm{cris}}/pA_{\mathrm{cris}}$. Recall that we remarked after the definition of $\Gamma(X, \mathcal{O}_X\widehat{\otimes}A_{\mathrm{cris}})$ that it equals the $p$-adic completion $C^+\widehat{\otimes}_{\mathbb{Z}_p}A_{\mathrm{cris}}$ of $C^+ \otimes_{\mathbb{Z}_p} A_{\mathrm{cris}}$. We start with the following:

*Claim.* For every element $c \in C^+\widehat{\otimes}_{\mathbb{Z}_p}A_{\mathrm{cris}}$ there are uniquely determined elements $a_i \in C^+$ for $i \in I$ such that for every $n \in \mathbb{N}$ the set $\{i \in I : a_i \notin p^nC^+\}$ is finite and $c = \sum_{i\in I} a_i \otimes b_i$ in $C^+\widehat{\otimes}_{\mathbb{Z}_p}A_{\mathrm{cris}}$.

To establish the claim one proves by induction that for every $n$ there are elements $a_{i,n} \in C^+$ for all $i \in I$, only finitely many of which are nonzero, such that $c - \sum_{i\in I} a_{i,n} \otimes b_i \in p^nC^+\widehat{\otimes}_{\mathbb{Z}_p}A_{\mathrm{cris}}$ and such that $a_{i,n} - a_{i,n-1} \in p^{n-1}C^+$. Namely, for $n = 0$ one can take $a_{i,0} = 0$ for all $i \in I$. In the induction step from $n$ to $n+1$ one considers an element $c' \in C^+\widehat{\otimes}_{\mathbb{Z}_p}A_{\mathrm{cris}}$ with $c - \sum_{i\in I} a_{i,n} \otimes b_i = p^nc'$. Then the image of $c'$ in $C^+\widehat{\otimes}_{\mathbb{Z}_p}A_{\mathrm{cris}}/(p) = C^+/pC^+ \otimes_{\mathbb{F}_p} A_{\mathrm{cris}}/pA_{\mathrm{cris}}$ can be written as $\sum_i \bar{\alpha}_i \otimes \bar{b}_i$ with uniquely determined elements $\bar{\alpha}_i \in C^+/pC^+$ which are zero for all but finitely many $i$. After choosing lifts $\alpha_i \in C^+$, the elements $a_{i,n+1} := a_{i,n} + p^n\alpha_i$ satisfy the assertion. Now taking $a_i$ as the limit of $a_{i,n}$ for $n \to \infty$ establishes the existence of the $a_i \in C^+$.

To prove the uniqueness, we must show that $\sum_{i\in I} a_i \otimes b_i = 0$ implies $a_i = 0$ for all $i$. It suffices to show that $a_i \in p^nC^+$ for all $n$ and $i$. This follows by induction on $n$, trivially starting with $n = 0$. If it holds for some $n$, we can write $a_i = p^na'_i$ for $a'_i \in C^+$. Then $p^n \cdot \sum_i a'_i \otimes b_i = \sum_i a_i \otimes b_i = 0$, and hence $\sum_i a'_i \otimes b_i = 0$, because $C^+\widehat{\otimes}_{\mathbb{Z}_p}A_{\mathrm{cris}}$ has no $p$-torsion by [Bourbaki 1961, Chapitre III, §5, no. 2, Théorème 1(v)] as $C^+$ and $A_{\mathrm{cris}}$ are flat over $\mathbb{Z}_p$. Considering the images $\bar{a}'_i$ of $a'_i$ in $C^+/pC^+$, the equation $\sum_i \bar{a}'_i \otimes \bar{b}_i = 0$ in $C^+\widehat{\otimes}_{\mathbb{Z}_p}A_{\mathrm{cris}}/(p) = C^+/pC^+ \otimes_{\mathbb{F}_p} A_{\mathrm{cris}}/pA_{\mathrm{cris}}$ implies that $\bar{a}'_i = 0$ in $C^+/pC^+$, whence $a'_i \in pC^+$ and $a_i \in p^{n+1}C^+$ as desired. This establishes our claim.

Furthermore we note that this claim (and in particular the uniqueness part) also applies if we replace $C^+$ by a finite free $\mathbb{Z}_p$-algebra (that is not necessarily reduced).

We lift $g$ to an element $\tilde{g} \in C\widehat{\otimes}_{\mathbb{Q}_p}A_{\mathrm{cris}}[1/p]$. After multiplying with a power of $p$ we can assume that $\tilde{g} \in C^+\widehat{\otimes}_{\mathbb{Z}_p}A_{\mathrm{cris}}$. By the claim we obtain uniquely determined elements $a_i \in C^+$ for all $i \in I$ with $\tilde{g} = \sum_i a_i \otimes b_i$ in $C^+\widehat{\otimes}_{\mathbb{Z}_p}A_{\mathrm{cris}}$. We show that $a_i \in \ker(C \to A)$ for all $i \neq 0$ which obviously implies $g \in \Gamma(Y, \mathcal{O}^+_Y) = A^+$.

As $C$ is noetherian the latter may be checked at completions $\widehat{C}_{\mathfrak{m}}$ of $C$ with respect to maximal ideals $\mathfrak{m}$ of $C$. If the point defined by $\mathfrak{m}$ is not in $\mathrm{Spa}(A, A^+) \subset \mathrm{Spa}(C, C^+)$ this claim is obvious. Otherwise we consider the surjections $C \twoheadrightarrow A \twoheadrightarrow A/\mathfrak{m}^nA = A'$ onto the finite-dimensional $\mathbb{Q}_p$-algebra $A'$, and let $A'^+$ denote the image of $C^+$ in $A'$. Then $A'^+$ is a finite $\mathbb{Z}_p$-algebra and we write $\bar{a}_i \in A'^+$ for the image of $a_i$. By what we noted above the expansion $\bar{g} = \sum \bar{a}_i \otimes b_i \in A'^+\widehat{\otimes}_{\mathbb{Z}_p}A_{\mathrm{cris}} = A'^+ \otimes_{\mathbb{Z}_p} A_{\mathrm{cris}}$ is unique and by assumption lies in $A'^+ \subset A'^+ \otimes_{\mathbb{Z}_p} A_{\mathrm{cris}}$. It follows that $\bar{a}_i = 0$ for all $i \neq 0$. We have shown that $a_i$ for $i \neq 0$ vanishes in $A' = A/\mathfrak{m}^n$ for all $n$ and the $a_i$ for $i \neq 0$ vanish in $A_{\mathfrak{m}}$.

(b), (c) We denote the residue field of $W$ by $k$ and let $k'$ be either $k$ for proving (c) or $\mathbb{F}_p$ for proving (b). We view the residue field $k(\!(u)\!)^{\mathrm{sep}} = \widetilde{\mathbb{A}}/p\widetilde{\mathbb{A}}$ of $\widetilde{\mathbb{A}}$ as a $k'(\!(u)\!)$-vector space. We denote the integral closure of $k'[\![u]\!]$ in $k(\!(u)\!)^{\mathrm{sep}}$ by $k[\![u]\!]^{\mathrm{sep}}$. It is a free $k'[\![u]\!]$-module: we can write $k(\!(u)\!)^{\mathrm{sep}}$ as union of finite extensions $E_i$ of $k'(\!(u)\!)$, where $E_i \subset E_{i+1}$, then $k[\![u]\!]^{\mathrm{sep}}$ is the increasing union of the rings of integers $\mathcal{O}_{E_i}$ which are free, and $\mathcal{O}_{E_i}$ is a direct summand of $\mathcal{O}_{E_{i+1}}$. Choosing the basis successively yields a basis for $k[\![u]\!]^{\mathrm{sep}}$.

We choose a $k'[\![u]\!]$-basis $(\bar{g}_i)_{i\in I}$ of $k[\![u]\!]^{\mathrm{sep}}$ with $\bar{g}_0 = 1$ and we lift the $\bar{g}_i$ to elements $g_i \in \widetilde{\mathbb{A}}^{[0,1)}$ with $g_0 = 1$.

We first prove (c) and use $k' = k$. The image of $g$ in

$$\Gamma(Y, \widetilde{\mathscr{A}}_Y^{[0,1)})/(p) = (A^+/pA^+ \otimes_{\mathbb{F}_p} k) \otimes_k k[\![u]\!]^{\mathrm{sep}}$$

can be written as $\sum_i \sum_{j=0}^{\infty} \bar{\alpha}_{i,j,0} \otimes u^j \bar{g}_i$ with uniquely determined elements $\bar{\alpha}_{i,j,0} \in A^+/pA^+ \otimes_{\mathbb{F}_p} k$ which are nonzero only for finitely many $i$ but possibly for all $j \geq 0$. After choosing lifts $\alpha_{i,j,0} \in A^+ \otimes_{\mathbb{Z}_p} W$, the image of $(1/p)\cdot(g - \sum_{i,j} \alpha_{i,j,0} \otimes u^j g_i)$ in $\Gamma(Y, \widetilde{\mathscr{A}}_Y^{[0,1)})/(p)$ can likewise be written as $\sum_{i,j} \bar{\alpha}_{i,j,1} \otimes u^j \bar{g}_i$ with uniquely determined elements $\bar{\alpha}_{i,j,1} \in A^+/pA^+ \otimes_{\mathbb{F}_p} k$. Note for this that $\Gamma(Y, \widetilde{\mathscr{A}}_Y^{[0,1)})$ has no $p$-torsion by [Bourbaki 1961, Chapitre III, §5, no. 2, Théorème 1(v)], because $\widetilde{\mathbb{A}}^{[0,1)}$ and $A^+$ are flat over $\mathbb{Z}_p$. Continuing in this way, we obtain elements $\alpha_{i,j} := \sum_{k=0}^{\infty} \alpha_{i,j,k}\, p^k \in A^+ \otimes_{\mathbb{Z}_p} W$ such that for every $n \geq 1$ the equality $g = \sum_{i,j} \alpha_{i,j} \otimes u^j g_i$ holds in $\Gamma(Y, \widetilde{\mathscr{A}}_Y^{[0,1)})/(p^n)$, although the sum does in general not converge in $\Gamma(Y, \widetilde{\mathscr{A}}_Y^{[0,1)})$.

The elements $\alpha_{i,j}$ are uniquely determined by $g$ because the equality $g = \sum_{i,j} \alpha_{i,j} \otimes u^j g_i$ in $\Gamma(Y, \widetilde{\mathscr{A}}_Y^{[0,1)})/(p^n)$ shows that the images of $\alpha_{i,j}$ in $A^+ \otimes_{\mathbb{Z}_p} W/(p^n)$ are uniquely determined for every $n$. The uniqueness of the $\alpha_{i,j}$ then follows from the fact that $A^+ \otimes_{\mathbb{Z}_p} W$ is $p$-adically separated. We conclude that the element $g$ lies in $\Gamma(Y, \mathscr{A}_Y^{[0,1)})$ if and only if $\alpha_{i,j} = 0$ whenever $i \neq 0$ or $j < 0$.

Now $g \otimes 1 \in \kappa(y)^+ \otimes_{\mathbb{Z}_p} \mathbb{A}^{[0,1)}$ implies that $\alpha_{i,j} \otimes 1 = 0$ in $\kappa(y)^+ \otimes_{\mathbb{Z}_p} W$ whenever $i \neq 0$ or $j < 0$. If this holds for every rigid analytic point $y$, then $\alpha_{i,j} = 0$ whenever $i \neq 0$ or $j < 0$, because $A^+ \otimes_{\mathbb{Z}_p} W$ is reduced. This implies $g \in \Gamma(Y, \mathscr{A}_Y^{[0,1)})$.

(b) Again the condition is necessary and we show that it is sufficient. Let us choose a closed immersion $Y = \mathrm{Spa}(A, A^+) \hookrightarrow X = \mathrm{Spa}(C, C^+)$ with $C$ a reduced Tate ring topologically of finite type over $\mathbb{Q}_p$. Then again our definitions imply that $\Gamma(Y, \widetilde{\mathscr{B}}_Y)$ is the quotient of $\Gamma(X, \widetilde{\mathscr{B}}_X)$ by the kernel of the epimorphism $C \twoheadrightarrow A$. We lift $g$ to an element $\tilde{g} \in C\widehat{\otimes}_{\mathbb{Q}_p}\widetilde{\mathbb{B}}$. After multiplying with a power of $p$ we can assume that $\tilde{g} \in C^+\widehat{\otimes}_{\mathbb{Z}_p}\widetilde{\mathbb{A}}$, where the complete tensor product denotes completion with respect to the $(p,u)$-adic topology.

We use the elements $g_i \in \widetilde{\mathbb{A}}^{[0,1)} \subset \widetilde{\mathbb{A}}$ with $g_0 = 1$ from the proof of (c) above (with $k' = \mathbb{F}_p$), whose residues $\bar{g}_i \in k[\![u]\!]^{\mathrm{sep}} \subset k(\!(u)\!)^{\mathrm{sep}}$ modulo $p$ form an $\mathbb{F}_p[\![u]\!]$-basis of $k[\![u]\!]^{\mathrm{sep}}$, and hence also an $\mathbb{F}_p(\!(u)\!)$-basis of $k(\!(u)\!)^{\mathrm{sep}}$. Then the image of $\tilde{g}$ in $C^+\widehat{\otimes}_{\mathbb{Z}_p}\widetilde{\mathbb{A}}/(p) = C^+/pC^+ \otimes_{\mathbb{F}_p} k[\![u]\!]^{\mathrm{sep}}$ can be written as $\sum_{i,j} \bar{\alpha}_{i,j,0} \otimes u^j \bar{g}_i$ with uniquely determined elements $\bar{\alpha}_{i,j,0} \in C^+/pC^+$ which are zero for all but finitely many $i$ and for $j \ll 0$. After choosing lifts $\alpha_{i,j,0} \in C^+$, the image of $(1/p)\cdot(g - \sum_{i,j} \alpha_{i,j,0} \otimes u^j g_i)$

in $C^+\widehat{\otimes}_{\mathbb{Z}_p}\widetilde{\mathbb{A}}/(p)$ can likewise be written as $\sum_{i,j}\bar{\alpha}_{i,j,1}\otimes u^j\bar{g}_i$ with uniquely determined elements $\bar{\alpha}_{i,j,1}\in C^+/pC^+$. Note for this that $C^+\widehat{\otimes}_{\mathbb{Z}_p}\widetilde{\mathbb{A}}$ has no $p$-torsion by [Bourbaki 1961, Chapitre III, §5, no. 2, Théorème 1(v)], because $\widetilde{\mathbb{A}}$ and $C^+$ are flat over $\mathbb{Z}_p$. Continuing in this way, we obtain elements $\alpha_{i,j}:=\sum_{k=0}^{\infty}\alpha_{i,j,k}\,p^k\in C^+$ such that for every $n\geq 1$ the equality $g=\sum_{i,j}\alpha_{i,j}\otimes u^j g_i$ holds in $C^+\widehat{\otimes}_{\mathbb{Z}_p}\widetilde{\mathbb{A}}/(p^n)$, although the sum does in general not converge in $C^+\widehat{\otimes}_{\mathbb{Z}_p}\widetilde{\mathbb{A}}$. The elements $\alpha_{i,j}$ are uniquely determined by $g$ by reasoning like in (c) above. We conclude that the element $g$ lies in $C^+$ if and only if $\alpha_{i,j}=0$ whenever $(i,j)\neq(0,0)$.

As $C$ is noetherian the latter may be checked at completions $\widehat{C}_{\mathfrak{m}}$ of $C$ with respect to maximal ideals $\mathfrak{m}$ of $C$. If the point defined by $\mathfrak{m}$ is not in $\mathrm{Spa}(A,A^+)\subset\mathrm{Spa}(C,C^+)$ this claim is obvious. Otherwise we consider the surjections $C\twoheadrightarrow A\twoheadrightarrow A/\mathfrak{m}^nA=A'$ onto the finite-dimensional $\mathbb{Q}_p$-algebra $A'$. Then our assumptions imply that the image of $a_{i,j}$ in $A'$ vanishes for $(i,j)\neq(0,0)$ by a similar reasoning as above for $A'^+$ in place of $C^+$. We have shown that the image of $a_{i,j}$ in $\widehat{C}_{\mathfrak{m}}$ lie in the kernel of $\widehat{C}_{\mathfrak{m}}\to\widehat{A}_{\mathfrak{m}}$ for all maximal ideals of $C$ and all $(i,j)\neq(0,0)$. The claim follows from this. □

**Remark 7.6.** Assume that in the situation of Lemma 7.5 the ring $A$ is reduced. We remark that it is then enough to check the conditions for surjections $A\twoheadrightarrow\kappa(y)$ for all rigid analytic points $y\in Y$. We only need to argue (in the situation of the proof above) that $g(y)\in\kappa(y)^+\subset\kappa(y)^+\widehat{\otimes}_{\mathbb{Z}_p}B$ implies that $a_i(y)=0$ in $\kappa(y)^+$ for all $i\neq 0$ for $B=A_{\mathrm{cris}}$, respectively $B=\widetilde{\mathbb{A}}$. Here we write $g(y)=1\otimes g\in\kappa(y)^+\widehat{\otimes}_{\mathbb{Z}_p}B$ and so on. If this holds for every rigid analytic point $y\in Y$, then $a_i=0$ for all $i\neq 0$, because $Y$ is reduced. This implies $g\in\Gamma(Y,\mathcal{O}_Y^+)$.

**Remark 7.7.** It is also possible to define $\mathbb{Z}$-filtrations $\mathrm{Fil}^i(\mathcal{O}_X\widehat{\otimes}B_{\mathrm{cris}})$ and $\mathrm{Fil}^i(\mathcal{O}_X\widehat{\otimes}B_{\mathrm{st}})$ on $\mathcal{O}_X\widehat{\otimes}B_{\mathrm{cris}}$ and $\mathcal{O}_X\widehat{\otimes}B_{\mathrm{st}}$, respectively. The most natural procedure seems to be the following: given $i\in\mathbb{Z}$ and an adic space $X=\mathrm{Spa}(A,A^+)$, a section $f\in\Gamma(X,\mathcal{O}_X\widehat{\otimes}B_{\mathrm{cris}})$ lies in $\Gamma(X,\mathrm{Fil}^i(\mathcal{O}_X\widehat{\otimes}B_{\mathrm{cris}}))$, if $f\otimes 1\in\mathrm{Fil}^i B_{\mathrm{cris}}\otimes_{\mathbb{Q}_p}B$ for all surjections $A\twoheadrightarrow B$ of A onto finite-dimensional $\mathbb{Q}_p$-algebras $B$. Here $\mathrm{Fil}^i B_{\mathrm{cris}}$ is the usual filtration on $B_{\mathrm{cris}}$ induced by restricting the $t$-adic filtration on Fontaine's ring $B_{\mathrm{dR}}$ to $B_{\mathrm{cris}}$. This construction obviously globalizes and defines a filtration of the sheaf $\mathcal{O}_X\widehat{\otimes}B_{\mathrm{cris}}$. A similar construction also applies to the filtration on $\mathcal{O}_X\widehat{\otimes}B_{\mathrm{st}}$. However some issues with this filtration seem to be a bit involved, in particular dealing with families. One main reason is, that $\mathcal{O}_X\widehat{\otimes}B^+_{\mathrm{cris}}$ is much better behaved than $\mathcal{O}_X\widehat{\otimes}B_{\mathrm{cris}}$, but $\mathrm{Fil}^0 B_{\mathrm{cris}}$ does not give back $B^+_{\mathrm{cris}}$. Hence we will not consider this filtration on $\mathcal{O}_X\widehat{\otimes}B_{\mathrm{cris}}$ explicitly.

**Proposition 7.8.** *Let $X$ be an adic space locally of finite type over $\mathbb{Q}_p$. The canonical inclusions induce equalities*

$$\widetilde{\mathscr{B}}_X^{\Phi=\mathrm{id}}=\mathcal{O}_X,\qquad\left(\mathcal{O}_X\widehat{\otimes}B^+_{\mathrm{cris}}\right)^{\Phi=\mathrm{id}}=\mathcal{O}_X,\qquad\left(\mathcal{O}_X\widehat{\otimes}B^+_{\mathrm{st}}\right)^{\Phi=\mathrm{id},N=0}=\mathcal{O}_X.$$

*Moreover one has*

$$(\mathcal{O}_X\widehat{\otimes}B^+_{\mathrm{cris}})^{\mathscr{G}_K}=(\mathcal{O}_X\widehat{\otimes}B^+_{\mathrm{st}})^{\mathscr{G}_K}=\mathcal{O}_X\otimes_{\mathbb{Q}_p}K_0.$$

*Proof.* It is clear that in all cases $\mathcal{O}_X$ injects onto the sheaves of invariants, and that $\mathcal{O}_X\otimes_{\mathbb{Q}_p}K_0$ injects into $(\mathcal{O}_X\widehat{\otimes}B_{\mathrm{cris}})^{\mathscr{G}_K}$. Let us prove the converse. Let $U=\mathrm{Spa}(A,A^+)\subset X$ be an affinoid open and

let $f \in \Gamma(U, \widetilde{\mathscr{B}}_X)$ be a section that is invariant under $\Phi$. Then for each quotient $A \twoheadrightarrow A'$ with $A'$ a finite-dimensional $\mathbb{Q}_p$-algebra the element $f \otimes 1 \in \Gamma(U, \widetilde{\mathscr{B}}_X) \otimes_A A' = A' \otimes_{\mathbb{Q}_p} \widetilde{\mathbb{B}}$ is invariant under $\Phi$ and hence $f \otimes 1 \in A' \subset A' \otimes_{\mathbb{Q}_p} \widetilde{\mathbb{B}}$. Now Lemma 7.5 implies $f \in \Gamma(U, \mathcal{O}_X)$. The other claims are proven using the same argument. □

**Definition 7.9.** Let $\mathscr{G}$ denote a compact topological group. A *family of $\mathscr{G}$-representations* on an adic space $X$ consists of a vector bundle $\mathcal{E}$ on $X$ together with an $\mathcal{O}_X$-linear action of the group $\mathscr{G}$ on $\mathcal{E}$ which is continuous for the topologies on the sections $\Gamma(-, \mathcal{E})$. This definition extends to the category of stacks on $\mathrm{Ad}^{\mathrm{lft}}_{\mathbb{Q}_p}$.

**Definition 7.10.** Let $X$ be an adic space locally of finite type over $\mathbb{Q}_p$.

(i) A *$\varphi$-module over $\mathscr{A}_X$* is an $\mathscr{A}_X$-module $M$ which is locally on $X$ free of finite rank over $\mathscr{A}_X$ together with an isomorphism $\Phi : \varphi^* M \xrightarrow{\sim} M$.

(ii) A *$\varphi$-module over $\mathscr{B}_X$* is an $\mathscr{B}_X$-module $M$ which is locally on $X$ free of finite rank over $\mathscr{B}_X$ together with an isomorphism $\Phi : \varphi^* M \xrightarrow{\sim} M$.

(iii) A $\varphi$-module $M$ over $\mathscr{B}_X$ is called *étale* if it is locally on $X$ of the form $N \otimes_{\mathscr{A}_X} \mathscr{B}_X$ for a $\varphi$-module $N$ over $\mathscr{A}_X$.

The following theorem summarizes results of [Hellmann 2013] which are needed in the sequel.

**Theorem 7.11.** *Let $X$ be a reduced adic space locally of finite type over $\mathbb{Q}_p$ and let $(\mathcal{N}, \Phi)$ be an étale $\varphi$-module of rank $d$ over $\mathscr{B}_X$.*

(i) *The set*

$$X^{\mathrm{adm}} = \{x \in X \mid \dim_{\kappa(x)}((\mathcal{N} \otimes_{\mathscr{B}_X} \widetilde{\mathscr{B}}_X) \otimes \kappa(x))^{\Phi=\mathrm{id}} = d\} \subset X$$

*is an open subspace and*

$$\mathcal{V} = (\mathcal{N} \otimes \widetilde{\mathscr{B}}_X)^{\Phi=\mathrm{id}}$$

*is a family of $\mathscr{G}_{K_\infty}$-representations on $X^{\mathrm{adm}}$.*

(ii) *If $f : Y \to X$ is a morphism in $\mathrm{Ad}^{\mathrm{lft}}$ and if $(\mathcal{N}_Y, \Phi_Y)$ denotes the pullback of $(\mathcal{N}, \Phi)$ along $f$, then $Y^{\mathrm{adm}} = f^{-1}(X^{\mathrm{adm}})$ and*

$$(\mathcal{N}_Y \otimes \widetilde{\mathscr{B}}_Y)^{\Phi=\mathrm{id}} = (f|_{Y^{\mathrm{adm}}})^* \mathcal{V}$$

*as families of $\mathscr{G}_{K_\infty}$-representations on $Y^{\mathrm{adm}}$.*

(iii) *If $(\mathfrak{M}, \Phi)$ is a $\varphi$-module of finite height over $\mathscr{A}_X^{[0,1)}$ as in Definition 6.1 and $(\mathcal{N}, \Phi) = (\mathfrak{M}, \Phi) \otimes_{\mathscr{A}_X^{[0,1)}} \mathscr{B}_X$, then*

$$U = X^{\mathrm{adm}} = \{x \in X \mid \mathrm{rk}_{\kappa(x)^+} \mathrm{Hom}_{\mathscr{A}_X^{[0,1)} \otimes \kappa(x), \Phi}(\mathfrak{M} \otimes \kappa(x), \widetilde{\mathscr{A}}_X^{[0,1)} \otimes \kappa(x)) = d\}$$

*and*

$$\mathscr{H}om_{\mathscr{A}_U^{[0,1)}, \Phi}(\mathfrak{M}|_U, \widetilde{\mathscr{A}}_U^{[0,1)}) \otimes_{\mathbb{Z}_p} \mathbb{Q}_p = \mathscr{H}om_{\mathscr{B}_U, \Phi}(\mathfrak{M}|_U \otimes_{\mathscr{A}_U^{[0,1)}} \mathscr{B}_U, \widetilde{\mathscr{B}}_U)$$

*as families of $\mathscr{G}_{K_\infty}$-representations on $U = X^{\mathrm{adm}}$.*

*Proof.* This is a summary of [Hellmann 2013, Propositions 8.20, 8.22, 8.23 and Corollary 8.21]. □

Given a cocharacter $\mu$ as in (2-5), the stack $\mathscr{H}_{\varphi,N,\preceq\mu}$ is the stack quotient of $P_{K_0,d} \times_{\operatorname{Spec}\mathbb{Q}_p} Q_{K,d,\mu}$ by the action of the reductive group $(\operatorname{Res}_{K_0/\mathbb{Q}_p} \mathrm{GL}_{d,K_0})_{E_\mu}$. Let us denote by $\mathscr{H}^{\mathrm{red}}_{\varphi,N,\preceq\mu}$ the quotient of the reduced subscheme underlying $P_{K_0,d} \times_{\operatorname{Spec}\mathbb{Q}_p} Q_{K,d,\preceq\mu}$ by the induced action of $(\operatorname{Res}_{K_0/\mathbb{Q}_p} \mathrm{GL}_{d,K_0})_{E_\mu}$. Recall that $P_{K_0,d} \times_{\operatorname{Spec}\mathbb{Q}_p} Q_{K,d,\mu}$ is reduced, hence this modification will not be necessary if we restrict to the case where the Hodge type is fixed by $\mu$.

**Corollary 7.12.** *There is an open substack $\mathscr{H}^{\mathrm{red,ad,adm}}_{\varphi,N,\preceq\mu} \subset \mathscr{H}^{\mathrm{red,ad,int}}_{\varphi,N,\preceq\mu}$ and a family $\mathcal{E}$ of $\mathscr{G}_{K_\infty}$-representations on $\mathscr{H}^{\mathrm{red,ad,adm}}_{\varphi,N,\preceq\mu}$ such that*

$$\mathcal{E} = (\underline{\mathcal{M}}(D, \Phi, N, \mathfrak{q}) \otimes_{\mathscr{B}^{[0,1)}_X} \widetilde{\mathscr{B}}_X)^{\Phi=\mathrm{id}},$$

*where $(D, \Phi, N, \mathfrak{q})$ denotes the restriction of the universal family on $\mathscr{H}^{\mathrm{red,ad}}_{\varphi,N,\preceq\mu}$. This subspace is maximal in the following sense: If $X$ is a reduced adic space and if $\underline{D}'$ is a $(\varphi, N)$-module with Hodge–Pink lattice over $X$ with Hodge polygon bounded by $\mu$, then the induced map $f : X \to \mathscr{H}^{\mathrm{ad}}_{\varphi,N,\preceq\mu}$ factors over $\mathscr{H}^{\mathrm{red,ad,adm}}_{\varphi,N,\preceq\mu}$ if and only if $X = X^{\mathrm{adm}}$ with respect to the family*

$$\underline{\mathcal{M}}(\underline{D}') \otimes_{\mathscr{B}^{[0,1)}_X} \widetilde{\mathscr{B}}.$$

*In this case there is a canonical isomorphism of $\mathscr{G}_{K_\infty}$-representations*

$$f^*\mathcal{E} = (\underline{\mathcal{M}}(\underline{D}') \otimes_{\mathscr{B}^{[0,1)}_X} \widetilde{\mathscr{B}})^{\Phi=\mathrm{id}}.$$

*If $L$ is a finite extension of $E_\mu$, then $\mathscr{H}^{\mathrm{red,ad,adm}}_{\varphi,N,\preceq\mu}(L) = \mathscr{H}^{\mathrm{red,ad,int}}_{\varphi,N,\preceq\mu}(L)$*

*Proof.* Let us write $X_{\preceq\mu} = (P_{K_0,d} \times_{\operatorname{Spec}\mathbb{Q}_p} Q_{K,d,\preceq\mu})^{\mathrm{red,ad}}$ for the moment. Further we denote the pullback of the universal family of vector bundles on the open unit disc to $X_{\preceq\mu}$ by $(\mathcal{M}, \Phi, N^{\mathcal{M}}_\nabla) = \underline{\mathcal{M}}(D, \Phi, N, \mathfrak{q})$. Locally on $X^{\mathrm{int}}_{\preceq\mu}$ there exists a $\varphi$-module of finite height $\mathfrak{M}$ inside $(\mathcal{M}, \Phi)$, at least after a Tate twist. It follows that $\mathfrak{M} \otimes_{\mathscr{A}^{[0,1)}_X} \mathscr{A}_X$ is étale and we may apply the above theorem. Then $X^{\mathrm{adm}}_{\preceq\mu} \subset X_{\preceq\mu}$ is invariant under the action of $(\operatorname{Res}_{K_0/\mathbb{Q}_p} \mathrm{GL}_{d,K_0})_{E_\mu}$ and hence its quotient by this group is an open substack $\mathscr{H}^{\mathrm{red,ad,adm}}_{\varphi,N,\preceq\mu} \subset \mathscr{H}^{\mathrm{red,ad}}_{\varphi,N,\preceq\mu}$. Further

$$(\mathcal{M} \otimes_{\mathscr{B}^{[0,1)}_{X^{\mathrm{adm}}_\mu}} \widetilde{\mathscr{B}}_{X^{\mathrm{adm}}_\mu})^{\Phi=\mathrm{id}}$$

is a $(\operatorname{Res}_{K_0/\mathbb{Q}_p} \mathrm{GL}_{d,K_0})_{E_\mu}$-equivariant vector bundle with $\mathscr{G}_{K_\infty}$-action on $X^{\mathrm{adm}}_{\preceq\mu}$. Hence it defines a family of $\mathscr{G}_{K_\infty}$-representations on $\mathscr{H}^{\mathrm{red,ad,adm}}_{\varphi,N,\preceq\mu}$.

The second statement is local on $X$ and hence, after locally choosing a basis of $D$, we can locally lift the morphism $f : X \to \mathscr{H}^{\mathrm{red,ad}}_{\varphi,N,\preceq\mu}$ to a morphism $f' : X \to X_{\preceq\mu}$ such that the pullback of $(D, \Phi, N, \mathfrak{q})$ on $X_{\preceq\mu}$ along $f'$ is isomorphic to $\underline{D}'$. Now the claim follows from Theorem 7.11 (ii). □

## 8. The universal semistable representation

In this section we want to construct a semistable $\mathscr{G}_K$-representation out of the $\mathscr{G}_{K_\infty}$-representation on $\mathscr{H}^{\mathrm{red,ad,adm}}_{\varphi,N,\preceq\mu}$ from Corollary 7.12. This will be possible only on a part of $\mathscr{H}^{\mathrm{red,ad,adm}}_{\varphi,N,\preceq\mu}$. First of all we need

to restrict to the open subspace where the Hodge polygon is constant. This can be seen as follows. Let $\mathcal{E}$ be a family of $\mathscr{G}_K$-representations on an adic space $X$. It follows from [Berger and Colmez 2008, §4.1] that the (generalized) Hodge–Tate weights vary continuously on $X$. Namely, they are the eigenvalues of Sen's operator $\Theta_{\mathrm{Sen}}$ constructed in [Berger and Colmez 2008, before Remark 4.1.3]. The characteristic polynomial of $\Theta_{\mathrm{Sen}}$ has coefficients in $\mathcal{O}_X \otimes_{\mathbb{Q}_p} K$. However, with any reasonable definition of a semistable family $\mathcal{E}$ the Hodge–Tate weights of $\mathcal{E} \otimes \kappa(x)$ should be integers for all $x \in X$ and hence the Hodge–Tate weights and the Hodge polygon are locally constant on $X$.

Secondly, Kisin [2006, Theorem 0.1 and Corollary 1.3.15] showed that the universal étale $(\varphi, N_\nabla)$-module $\underline{\mathcal{M}}$ on $\mathscr{H}^{\mathrm{ad,int}}_{\varphi,N,\preceq\mu}$ from Corollary 6.6 can come from a semistable $\mathscr{G}_K$-representation only if the connection $\nabla$ has logarithmic singularities, which is equivalent to $N_\nabla^{\mathcal{M}}$ being holomorphic; see Remark 4.2 (2). Therefore we have to restrict further to the closed subspace $\mathscr{H}^{\nabla}_{\varphi,N,\mu} \cap \mathscr{H}^{\mathrm{ad,adm}}_{\varphi,N,\mu}$ of $\mathscr{H}^{\mathrm{ad,adm}}_{\varphi,N,\mu}$ which is isomorphic to $\mathscr{D}^{\mathrm{ad,adm}}_{\varphi,N,\mu}$. Here $\mathscr{D}^{\mathrm{ad,adm}}_{\varphi,N,\mu} \subset \mathscr{D}^{\mathrm{ad}}_{\varphi,N,\mu}$ is the admissible locus with respect to the family defined in Remark 4.10.

**Lemma 8.1.** *Let $\mathcal{E}$ be a family of $\mathscr{G}_{K_\infty}$-representations over a reduced adic space $X$ locally of finite type over $\mathbb{Q}_p$. Let $\mathfrak{M}_1$ and $\mathfrak{M}_2$ be two $\varphi$-modules of finite height over $\mathscr{A}_X^{[0,1)}$ such that $\varphi$ and $\mathscr{G}_{K_\infty}$-equivariant isomorphisms*

$$\mathfrak{M}_i \otimes_{\mathscr{A}_X^{[0,1)}} \widetilde{\mathscr{A}}_X^{[0,1)}[1/p] \cong \mathcal{E} \otimes_{\mathcal{O}_X} \widetilde{\mathscr{A}}_X^{[0,1)}[1/p] \tag{8-1}$$

*exist for $i = 1, 2$. Then $\mathfrak{M}_1[1/p] = \mathfrak{M}_2[1/p]$ as $\mathscr{A}_X^{[0,1)}[1/p]$-submodules of $\mathcal{E} \otimes_{\mathcal{O}_X} \widetilde{\mathscr{A}}_X^{[0,1)}[1/p]$. In particular they are isomorphic as $\varphi$-modules.*

*Proof.* The case $X = \operatorname{Spa}\mathbb{Q}_p$ was proven by Kisin [2006, Proposition 2.1.12].

If $X = \operatorname{Spa}(A, A^+)$ for a finite free $\mathbb{Z}_p$-algebra $A^+$ and $A = A^+[1/p]$, then this implies that $\mathfrak{M}_1[1/p]$ and $\mathfrak{M}_2[1/p]$ agree as $\mathscr{A}^{[0,1)}_{\operatorname{Spa}(\mathbb{Q}_p,\mathbb{Z}_p)}[1/p]$-submodules (even without the $A$-action).

For general $X$ we may work locally and assume that $\mathfrak{M}_i \cong (\mathscr{A}_X^{[0,1)})^n$. The isomorphisms (8-1) yield a matrix $M \in \mathrm{GL}_n(\Gamma(X, \widetilde{\mathscr{A}}_X^{[0,1)}[1/p]))$ and we must show that $M \in \mathrm{GL}_n(\Gamma(X, \mathscr{A}_X^{[0,1)}[1/p]))$. It suffices to show that every entry $g$ of $M$ and $M^{-1}$ lies in $\Gamma(X, \mathscr{A}_X^{[0,1)}[1/p])$. Multiplying the entry $g$ by a power of $p$ we can assume that it lies in $\Gamma(X, \widetilde{\mathscr{A}}_X^{[0,1)})$. By Lemma 7.5 (c) we must check that $g(x) \in \kappa(x)^+ \widehat{\otimes}_{\mathbb{Z}_p} \mathbb{A}^{[0,1)}$ for every rigid analytic point $x \in X$. Since $\kappa(x)$ is a finite-dimensional $\mathbb{Q}_p$-algebra, this was proved above. □

**Definition 8.2.** Let $\mathcal{E}$ be a family of $\mathscr{G}_K$-representations of rank $d$ on an adic space $X$ locally of finite type over $\mathbb{Q}_p$. Denote by $\overline{X}$ the reduced subspace underlying $X$ and by $\bar{\mathcal{E}}$ the restriction of $\mathcal{E}$ to $\overline{X}$.

(i) The family $\mathcal{E}$ is said to be *crystalline with negative Hodge Tate weights* if fpqc-locally on $X$ there is a $\varphi$-module $\overline{\mathfrak{M}}$ of finite height over $\mathscr{A}_{\overline{X}}^{[0,1)}$ and a $\varphi$ and $\mathscr{G}_{K_\infty}$-equivariant isomorphism

$$\overline{\mathfrak{M}} \otimes_{\mathscr{A}_{\overline{X}}^{[0,1)}} \widetilde{\mathscr{A}}_{\overline{X}}^{[0,1)}\left[\frac{1}{p}\right] \cong \bar{\mathcal{E}} \otimes_{\mathcal{O}_{\overline{X}}} \widetilde{\mathscr{A}}_{\overline{X}}^{[0,1)}\left[\frac{1}{p}\right] \tag{8-2}$$

and a $(\varphi, N_\nabla)$-module $\mathcal{M}$ over $\mathscr{B}_X^{[0,1)}$ deforming $\overline{\mathfrak{M}} \otimes_{\mathscr{A}_{\overline{X}}^{[0,1)}} \mathscr{B}_{\overline{X}}^{[0,1)}$ as a $\varphi$-module such that (8-2) extends to a $(\mathscr{G}_K, \varphi)$ equivariant isomorphism

$$\mathcal{M} \otimes_{\mathscr{B}_X^{[0,1)}} B^+_{\mathrm{cris}} \widehat{\otimes} \mathcal{O}_X \cong \mathcal{E} \otimes_{\mathcal{O}_X} B^+_{\mathrm{cris}} \widehat{\otimes} \mathcal{O}_X.$$

(ii) The family $\mathcal{E}$ is said to be *semistable with negative Hodge Tate weights* if fpqc-locally on $X$ there is a $\varphi$-module $\overline{\mathfrak{M}}$ of finite height over $\mathscr{A}_{\overline{X}}^{[0,1)}$ and a $\varphi$ and $\mathscr{G}_{K_\infty}$-equivariant isomorphism

$$\overline{\mathfrak{M}} \otimes_{\mathscr{A}_{\overline{X}}^{[0,1)}} \widetilde{\mathscr{A}}_{\overline{X}}^{[0,1)}\Big[\frac{1}{p}\Big] \cong \bar{\mathcal{E}} \otimes_{\mathcal{O}_{\overline{X}}} \widetilde{\mathscr{A}}_{\overline{X}}^{[0,1)}\Big[\frac{1}{p}\Big] \tag{8-3}$$

and a $(\varphi, N_\nabla)$-module $\mathcal{M}$ over $\mathscr{B}_X^{[0,1)}$ deforming $\overline{\mathfrak{M}} \otimes_{\mathscr{A}_{\overline{X}}^{[0,1)}} \mathscr{B}_{\overline{X}}^{[0,1)}$ as a $\varphi$-module such that (8-3) extends to a $(\mathscr{G}_K, \varphi, N)$ equivariant isomorphism

$$\mathcal{M} \otimes_{\mathscr{B}_X^{[0,1)}} B_{\mathrm{st}}^+ \widehat{\otimes} \mathcal{O}_X \cong \mathcal{E} \otimes_{\mathcal{O}_X} B_{\mathrm{st}}^+ \widehat{\otimes} \mathcal{O}_X. \tag{8-4}$$

(iii) We say that $\mathcal{E}$ is *crystalline* (resp. *semistable*) if some twist of $\mathcal{E}$ with a power of the cyclotomic character is crystalline with negative Hodge–Tate weights (resp. semistable with negative Hodge–Tate weights).

**Remark 8.3.** The definition of being crystalline or semistable is slightly involved. We did not define it in the usual way using only period ring $B_{\mathrm{cris}}^+ \widehat{\otimes} \mathcal{O}_X$, as our method requires that we have a comparison isomorphism for the integral models on the open unit disc as in (8-2). Working only with $B_{\mathrm{cris}}^+ \widehat{\otimes} \mathcal{O}_X$ it is not clear to us whether this is automatically true.

**Lemma 8.4.** *Let $X$ be an adic space locally of finite type over $\mathbb{Q}_p$ and let $\mathcal{E}$ be a family of crystalline (resp. semistable) $\mathscr{G}_K$-representations with negative Hodge–Tate weights on $X$. Assume that the objects $\overline{\mathfrak{M}}$ and $\mathcal{M}$ in the above definition exist globally on $X$. Then the following holds true*:

(i) *If $X = \mathrm{Spa}(A, A^+)$ for some finite-dimensional $\mathbb{Q}_p$-algebra $A$, then $\Gamma(X, \mathcal{E})$ is crystalline (resp. semistable) as a $\mathscr{G}_K$-representation on a finite dimensional $\mathbb{Q}_p$-vector space and $\underline{D}(\mathcal{M}) = D_{\mathrm{cris}}(\Gamma(X, \mathcal{E}))$ (resp. $= D_{\mathrm{st}}(\Gamma(X, \mathcal{E}))$) as filtered $\varphi$-modules (resp. as filtered $(\varphi, N)$-modules), compatible with the canonical $A$-action on both sides.*

(ii) *If $Y \to X$ is any morphism of adic spaces locally of finite type, and if $\mathcal{E}_Y$ denotes the $\mathscr{G}_K$-representation on $Y$ obtained by base changing $\mathcal{E}$, then $\mathcal{E}_Y$ is crystalline (resp. semistable).*

(iii) *Assume that $\mathcal{E}$ is crystalline (resp. semistable) with negative Hodge–Tate weights. The family $\mathcal{M}$ is uniquely determined as a $(\varphi, N_\nabla)$-module and in fact as a submodule of $\mathcal{E} \otimes_{\mathcal{O}_X} B_{\mathrm{cris}}^+ \widehat{\otimes} \mathcal{O}_X$ (resp. of $\mathcal{E} \otimes_{\mathcal{O}_X} B_{\mathrm{st}}^+ \widehat{\otimes} \mathcal{O}_X$)*

*Proof.* (i) This follows from (the covariant formulation of) [Kisin 2006, Proposition 2.1.5], the proof of which implies that the morphisms in [loc. cit., (2.1.6)] are isomorphisms. The fact that the morphism is compatible with the $A$-action follows from functoriality.

(ii) This is obvious.

(iii) We only prove the crystalline case. The semi-stable case is proved along the same lines. Assume that $X = \mathrm{Spa}(A, A^+)$ is affinoid and that there are two $\mathscr{B}_X^{[0,1)}$-modules $\mathcal{M}_1$ and $\mathcal{M}_2$ as in the definition. We set $D_i = \underline{D}(\mathcal{M}_i)$ and consider the morphisms

$$D_i \to D_i \otimes_{\mathcal{O}_X} B_{\mathrm{cris}}^+ \widehat{\otimes} \mathcal{O}_X \to \mathcal{M}_i \otimes_{\mathscr{B}_X^{[0,1)}} B_{\mathrm{cris}}^+ \widehat{\otimes} \mathcal{O}_X = \mathcal{E} \otimes_{\mathcal{O}_X} B_{\mathrm{cris}}^+ \widehat{\otimes} \mathcal{O}_X.$$

As these morphisms are compatible with the $\mathscr{G}_K$-action (which is of course trivial on $D_i$) we obtain a morphism

$$\alpha_i : D_i \to (\mathcal{E} \otimes_{\mathcal{O}_X} B^+_{\mathrm{cris}} \widehat{\otimes} \mathcal{O}_X)^{\mathscr{G}_K}.$$

Now both sides are locally on $X$ free as $\mathcal{O}_X \otimes_{\mathbb{Q}_p} K_0$-modules. To see this on the right-hand side use the equality

$$\mathcal{M}_i \otimes_{\mathscr{B}_X^{[0,1)}} B^+_{\mathrm{cris}} \widehat{\otimes} \mathcal{O}_X = \mathcal{E} \otimes_{\mathcal{O}_X} B^+_{\mathrm{cris}} \widehat{\otimes} \mathcal{O}_X$$

and apply Proposition 7.8. Now the construction of this map is functorial and for each quotient $A \twoheadrightarrow A'$ onto a finite-dimensional $\mathbb{Q}_p$-algebra $A'$ the induced map

$$\alpha_{i,A'} : D_i \otimes_A A' \to (\mathcal{E} \otimes_{\mathcal{O}_X} B^+_{\mathrm{cris}} \widehat{\otimes} \mathcal{O}_X)^{\mathscr{G}_K} \otimes_A A' = ((\mathcal{E} \otimes_A A') \otimes_{\mathcal{O}_X} B^+_{\mathrm{cris}} \widehat{\otimes} \mathcal{O}_X)^{\mathscr{G}_K}.$$

is an isomorphism. It follows that $\alpha$ is an isomorphism for $i = 1, 2$ and hence $D = D_1 = D_2$ is uniquely determined as a $\varphi$-submodule of $\mathcal{E} \otimes_{\mathcal{O}_X} B^+_{\mathrm{cris}} \widehat{\otimes} \mathcal{O}_X$. In particular we have shown that $\mathcal{M}_i[1/\lambda]$ is uniquely determined as a submodule of $\mathcal{E} \otimes_{\mathcal{O}_X} B_{\mathrm{cris}} \widehat{\otimes} \mathcal{O}_X$.

It remains to prove that the two filtrations on $D = D_1 = D_2$ are the same. Assume this is not the case. Then there exists a surjection $A \twoheadrightarrow A'$ onto a finite-dimensional $\mathbb{Q}_p$-algebra $A'$ such that the filtrations on $D \otimes_A A'$ induced by $D_1$ and $D_2$ do not agree. Replacing $A$ by $A'$ we may assume that $A'$ is a finite dimensional $\mathbb{Q}_p$-algebra. However, in this case (i) implies that $\mathcal{M}_1 = \mathcal{M}_2$ (as submodules of $\mathcal{E} \otimes_{\mathbb{Q}_p} B^+_{\mathrm{cris}}$) and hence the filtrations on $D_1$ and $D_2$ coincide. □

**Remark 8.5.** Let $\mathcal{E}$ be a crystalline representation with negative Hodge–Tate weights. Then fpqc-locally on $X$ we have associated a $(\varphi, N_\nabla)$-module $\mathcal{M}$ over $\mathscr{B}_X^{[0,1)}$ as in Definition 8.2. By the uniqueness result established in the previous lemma and fpqc descent this $(\varphi, N_\nabla)$-module in fact descends to $X$. The same remark applies to semistable representations as well.

Using this remark we can make the following definition:

**Definition 8.6.** Let $X$ be an adic space locally of finite type over $\mathbb{Q}_p$ and let $\mathcal{E}$ be a family of $\mathscr{G}_K$-representations on $X$.

(i) Assume that $\mathcal{E}$ is crystalline with negative Hodge–Tate weights and let $\mathcal{M}$ as in Definition 8.2. Then define $D_{\mathrm{cris}}(\mathcal{E}) = \underline{D}(\mathcal{M})$.

(ii) Assume that $\mathcal{E}$ is semistable with negative Hodge–Tate weights and let $\mathcal{M}$ as in Definition 8.2. Then define $D_{\mathrm{st}}(\mathcal{E}) = \underline{D}(\mathcal{M})$.

(iii) Assume that $\mathcal{E}$ is crystalline and that its twist $\mathcal{E}(i)$ is crystalline with negative Hodge–Tate weights for some $i \in \mathbb{Z}$. Then define $D_{\mathrm{cris}}(\mathcal{E}) = D_{\mathrm{cris}}(\mathcal{E}(i))(-i)$.

(iv) Assume that $\mathcal{E}$ is semistable and that its twist $\mathcal{E}(i)$ is semistable with negative Hodge–Tate weights for some $i \in \mathbb{Z}$. Then define $D_{\mathrm{st}}(\mathcal{E}) = D_{\mathrm{st}}(\mathcal{E}(i))(-i)$.

**Remark 8.7.** Obviously the last two parts of the definition are independent of the choice of $i$ such that $\mathcal{E}(i)$ has negative Hodge–Tate weights.

The above defines a functor from the category of crystalline representations on $X$ to the category of filtered $\varphi$-modules over $X$. Moreover it is a direct consequence of the definition that for every morphism $f : Y \to X$ and any family of crystalline $\mathscr{G}_K$-representations on $X$ we have

$$D_{\mathrm{cris}}(f^*\mathcal{E}) = f^* D_{\mathrm{cris}}(\mathcal{E}).$$

The same remark applies to the semistable case as well.

**Definition 8.8.** Let $\mu$ be a cocharacter as in (2-5), let $E_\mu$ be its reflex field, and let $X$ be an adic space locally of finite type over $E_\mu$. We say that a crystalline (resp. semistable) $\mathscr{G}_K$-representation $\mathcal{E}$ over $X$ has *constant Hodge polygon equal to* $\mu$ if the $K$-filtered $\varphi$-module $D_{\mathrm{cris}}(\mathcal{E})$ (resp. $D_{\mathrm{st}}(\mathcal{E})$) over $X$ has this property.

It is obvious from the definition that $D_{\mathrm{st}}$ defines a functor from the category of semistable representations with constant Hodge polygon $\mu$ over an adic space $X$ to the category of $K$-filtered $(\varphi, N)$-modules over $X$ with constant Hodge polygon $\mu$ and similarly for crystalline representations.

**Remark 8.9.** Let $\mathcal{E}$ be a crystalline (resp. semistable) representation over $X$ with negative Hodge–Tate weights and let $\mathcal{M}$ be as in Definition 8.2. We write $D = D_{\mathrm{cris}}(\mathcal{E}) = \underline{D}(\mathcal{M})$. Then

$$D \otimes_{\mathcal{O}_X \otimes_{\mathbb{Q}_p} K_0} \mathscr{B}_X^{[0,1)}[1/\lambda] \cong \mathcal{M} \otimes_{\mathscr{B}_X^{[0,1)}} \mathscr{B}_X^{[0,1)}[1/\lambda]$$

and hence, as $\lambda$ is invertible in $B_{\mathrm{cris}}$, we obtain a $(\mathscr{G}_K, \varphi)$-equivariant isomorphism

$$D_{\mathrm{cris}}(\mathcal{E}) \otimes_{\mathcal{O}_X \otimes_{\mathbb{Q}_p} K_0} \mathcal{O}_X \widehat{\otimes} B_{\mathrm{cris}} \cong \mathcal{E} \otimes_{\mathcal{O}_X} \mathcal{O}_X \widehat{\otimes} B_{\mathrm{cris}}.$$

Similarly, if $\mathcal{E}$ is semistable, we obtain a $(\mathscr{G}_K, \varphi, N)$-equivariant isomorphism

$$D_{\mathrm{st}}(\mathcal{E}) \otimes_{\mathcal{O}_X \otimes_{\mathbb{Q}_p} K_0} \mathcal{O}_X \widehat{\otimes} B_{\mathrm{st}} \cong \mathcal{E} \otimes_{\mathcal{O}_X} \mathcal{O}_X \widehat{\otimes} B_{\mathrm{st}}.$$

Using twists by the cyclotomic character, we find that the same holds true also for crystalline (resp. semistable) representations with arbitrary Hodge–Tate weights.

Moreover, if we had defined (the correct) filtration on $\mathcal{O}_X \widehat{\otimes} B_{\mathrm{st}}$ these morphisms would also respect filtrations. However, as we will not explicitly make use of this, we did not carefully define the filtrations.

**Lemma 8.10.** *Let $\mathcal{E}$ be a family of $\mathcal{G}_K$-representations on an adic space $X$ locally of finite type over $\mathbb{Q}_p$. Then $\mathcal{E}$ is crystalline if and only it is semistable and the monodromy $N$ on $D_{\mathrm{st}}(\mathcal{E})$ vanishes. In this case we have $D_{\mathrm{st}}(\mathcal{E}) = D_{\mathrm{cris}}(\mathcal{E})$ as subobjects of $\mathcal{E} \otimes_{\mathcal{O}_X} (\mathcal{O}_X \widehat{\otimes} B_{\mathrm{st}})$.*

*Proof.* We may assume that $\mathcal{E}$ has negative Hodge–Tate weights and that there exists some $\mathcal{M}$ as in Definition 8.2.

Assume that $\mathcal{E}$ is semistable with vanishing monodromy. As the isomorphism (8-4) is equivariant for the action of $N$ the claim follows after taking $N = 0$ on both sides.

Conversely, let us assume that $\mathcal{E}$ is crystalline. Then obviously $\mathcal{E}$ is semistable and using the definition of $D_{\mathrm{st}}$ we see immediately that $N = 0$ on $D_{\mathrm{st}}(\mathcal{E})$. □

**Remark 8.11.** In [Kisin 2008], techniques from [Kisin 2006] are used to construct what are called *(potentially) semistable deformation rings*. Fix a continuous representation $\bar{\rho} : \mathcal{G}_K \to \mathrm{GL}_n(\mathbb{F})$ with $\mathbb{F}$ a finite extension of $\mathbb{F}_p$ as well as a set of labeled Hodge–Tate weights

$$\mathbf{k} = \{k_{i,\tau}, i = 1, \ldots, n, \tau : K \hookrightarrow \overline{\mathbb{Q}}_p\}.$$

Given these data, Kisin constructs a quotient $R_{\bar{\rho}}^{\mathbf{k}}$ of the universal framed deformation ring[1] $R_{\bar{\rho}}$ of $\bar{\rho}$ such that a point

$$\operatorname{Spec} L \to \operatorname{Spec} R_{\bar{\rho}}$$

with $L$ a finite extension of $\mathbb{Q}_p$, factors over $\operatorname{Spec} R_{\bar{\rho}}^{\mathbf{k}}$ if and only if the corresponding Galois representation is semistable[2] with labeled Hodge–Tate weights $\mathbf{k}$. As the defining condition for $R_{\bar{\rho}}^{\mathbf{k}}$ is only formulated for points, the ring $R_{\bar{\rho}}^{\mathbf{k}}$ is reduced by definition (once it is known to exist). Kisin moreover shows that for every finite-dimensional $\mathbb{Q}_p$-algebra $A$ a morphism

$$\operatorname{Spec} A \to \operatorname{Spec} R_{\bar{\rho}}$$

factors over $\operatorname{Spec} R_{\bar{\rho}}^{\mathbf{k}}$ if and only if the corresponding representation $\rho : \mathcal{G}_K \to \mathrm{GL}_n(A)$ is semistable.

Our construction differs from Kisin's strategy in the following way: Kisin starts with a family of Galois representations on *integral* level and cuts out the locus in the generic fiber where the representations are semistable. In contrast to this we start with a family of $p$-adic Hodge structures in characteristic zero and cut out the locus where this family of $p$-adic Hodge structures comes from a Galois representation.

On the other hand after having constructed a universal family in our case, we can compare the outcome of this construction to Kisin's deformation space again. This is done in Proposition 8.17 below.

**Lemma 8.12.** *Let $X$ be an adic space locally of finite type over $\mathbb{Q}_p$ and let $\mathcal{E}$, $\mathcal{E}_1$ and $\mathcal{E}_2$ be families of semistable representations.*

(i) *Assume that $\mathcal{E}$ has negative Hodge–Tate weights. Then there is a canonical isomorphism*

$$\mathcal{E} \to (\underline{\mathcal{M}}(D_{\mathrm{st}}(\mathcal{E})) \otimes_{\mathscr{B}_X^{[0,1)}} (\mathcal{O}_X \widehat{\otimes} B_{\mathrm{st}}^+))^{\varphi = \mathrm{id}, N = 0}.$$

(ii) *One has $\mathcal{E}_1 \cong \mathcal{E}_2$ if and only if $D_{\mathrm{st}}(\mathcal{E}_1) \cong D_{\mathrm{st}}(\mathcal{E}_2)$.*

*Proof.* (i) Let us write $\mathcal{M} = \mathcal{M}(D_{\mathrm{st}}(\mathcal{E}))$. Then by definition fpqc-locally on $X$ we obtain an isomorphism

$$\mathcal{M} \otimes_{\mathscr{B}_X^{[0,1)}} B_{\mathrm{st}}^+ \widehat{\otimes} \mathcal{O}_X \cong \mathcal{E} \otimes_{\mathcal{O}_X} B_{\mathrm{st}}^+ \widehat{\otimes} \mathcal{O}_X.$$

Then locally on $X$ the claim follows by applying the invariants on both sides and using Proposition 7.8. The construction of this morphism is obviously compatible with the descent data and hence descends to $X$.

(ii) After twisting with powers of the cyclotomic character, we may assume that $\mathcal{E}_1$ and $\mathcal{E}_2$ have negative Hodge–Tate weights. Then second part is a direct consequence of the first. □

[1] Note that our notations here differ from Kisin's.

[2] There is a similar version with *crystalline* instead of *semistable*.

**Proposition 8.13.** *Let $X$ be a reduced adic space locally of finite type over $E_\mu$ and let $\underline{D}=(D,\Phi,N,\mathcal{F}^\bullet)\in \mathscr{D}^{\mathrm{an,adm}}_{\varphi,N,\mu}(X)$. Then there is a family of semistable representations $\mathcal{E}$ on $X$ such that $D_{\mathrm{st}}(\mathcal{E})=(D,\Phi,N,\mathcal{F}^\bullet)$. Moreover, $\mathcal{E}$ is canonically identified with the subrepresentation*

$$(\underline{\mathcal{M}}(D_{\mathrm{st}}(\mathcal{E}))\otimes_{\mathscr{B}_X^{[0,1)}}(\mathcal{O}_X\widehat{\otimes}B^+_{\mathrm{st}}))^{\varphi=\mathrm{id},N=0}$$

*of $\underline{\mathcal{M}}(D_{\mathrm{st}}(\mathcal{E}))\otimes_{\mathscr{B}_X^{[0,1)}}(\mathcal{O}_X\widehat{\otimes}B^+_{\mathrm{st}})$.*

*Proof.* After twisting with powers of the cyclotomic character and after changing $\mu$ accordingly, we may assume that the Hodge–Tate weights defined by $\mu$ are negative. Let us write $\mathfrak{M}$ for the choice of a $\varphi$-module of finite height over $\mathscr{A}_X^{[0,1)}$ and a family $\mathcal{E}$ of $\mathscr{G}_{K_\infty}$-representations such that there is a $(\varphi,\mathscr{G}_{K_\infty})$-equivariant isomorphism

$$\mathcal{E}\otimes_{\mathcal{O}_X}\widetilde{\mathscr{A}}_X^{[0,1)}[1/p]\cong\mathfrak{M}\otimes_{\mathscr{A}_X^{[0,1)}}\widetilde{\mathscr{A}}_X^{[0,1)}.$$

Such a module exists fpqc locally on $X$ by definition of the admissible locus and Theorem 7.11 (iii).

This isomorphism extends to an isomorphism

$$\mathcal{E}\otimes_{\mathcal{O}_X}(\mathcal{O}_X\widehat{\otimes}B^+_{\mathrm{st}})\cong\mathcal{M}\otimes_{\mathscr{B}_X^{[0,1)}}(\mathcal{O}_X\widehat{\otimes}B^+_{\mathrm{st}})$$

that is still equivariant for the actions of $\varphi$ and $\mathscr{G}_{K_\infty}$. According to the definition of a semistable representation we have to prove that the $\mathscr{G}_{K_\infty}$ action on $\mathcal{E}$ extends to an action of $\mathscr{G}_K$ and that the above isomorphism is equivariant for $\mathscr{G}_K$. As $\mathcal{E}$ embeds into the left-hand side, it is enough to show that it is stabilized by the $\mathscr{G}_K$-action on the right-hand side. After localization we may assume that $X=\mathrm{Spa}(A,A^+)$ is affinoid and that $\mathcal{E}$ is the trivial vector bundle on $X$. After choosing a basis of $\mathcal{E}$ let $g\in\mathscr{G}_K$ and denote by $M\in\mathrm{Mat}_{n\times n}(\Gamma(X,\mathcal{O}_X\widehat{\otimes}B^+_{\mathrm{st}}))$ the matrix of the $g$-action with respect to this basis. We have to show that this matrix has entries in $A$. However, $M\otimes_A\kappa(x)$ has entries in $\kappa(x)$ for all classical points $x\in X$ by [Kisin 2006, Proposition 2.1.5] (note that the proof of that proposition implies that the arrows in (2.1.6) of [loc. cit.] are isomorphisms). It now follows from Lemma 7.5 (a) (and the remark following that lemma) that $M$ has entries in $A$.

This proves the existence of $\mathcal{E}$ fpqc-locally on $X$. In order to finish the proof, we just notice that our construction defines descend data on $\mathcal{E}$ that are compatible with the isomorphisms

$$D_{\mathrm{st}}(\mathcal{E})\to(D,\Phi,N,\mathcal{F}^\bullet)$$

and the descend data on the latter. Hence both $\mathcal{E}$ as well as the isomorphism descent. □

Recall that the stack $\mathscr{D}^{\mathrm{ad}}_{\varphi,N,\mu}$ is the quotient of the adic space $X_\mu$ associated to $P_{K_0,d}\times\mathrm{Flag}_{K,d,\mu}$ by the action of the group $(\mathrm{Res}_{K_0/\mathbb{Q}_p}\mathrm{GL}_{d,K_0})_{E_\mu}$ and consider the open subspace $X_\mu^{\mathrm{adm}}\subset X_\mu^{\mathrm{int}}\subset X_\mu$. This subset is stable under the action of $(\mathrm{Res}_{K_0/\mathbb{Q}_p}\mathrm{GL}_{d,K_0})_{E_\mu}$ and we write $\mathscr{D}^{\mathrm{ad,adm}}_{\varphi,N,\mu}$ for the quotient of $X_\mu^{\mathrm{adm}}$ by this action.

**Proposition 8.14.** *Let $\mu$ be a cocharacter as in* (2-5) *and let $E_\mu$ be its reflex field. Let $X$ be a reduced adic space locally of finite type over $E_\mu$ and let $\mathcal{E}$ be a family of semistable $\mathscr{G}_K$-representations on $X$ with*

*constant Hodge polygon equal to $\mu$. Then the morphism*

$$X \to \mathscr{D}^{\mathrm{ad}}_{\varphi,N,\mu}$$

*induced by the $K$-filtered $(\varphi, N)$-module $D_{\mathrm{st}}(\mathcal{E})$ factors over $\mathscr{D}^{\mathrm{ad,adm}}_{\varphi,N,\mu}$.*

*Proof.* By Definition 8.2, there is (locally on $X$) an $\mathscr{A}_X^{[0,1)}$-module $\mathfrak{M}$ such that $\mathfrak{M} \otimes_{\mathscr{A}_X^{[0,1)}} \mathscr{B}_X^{[0,1)} = \underline{\mathcal{M}}(D)$, where $D = D_{\mathrm{st}}(\mathcal{E})$ is the filtered $(\varphi, N)$-module on $X$ defining the morphism $X \to \mathscr{D}^{\mathrm{ad}}_{\varphi,N,\mu}$. Moreover by definition

$$\mathfrak{M} \otimes_{\mathscr{A}_X^{[0,1)}} \widetilde{\mathscr{A}}_X^{[0,1)}[1/p] \cong \mathcal{E} \otimes_{\mathcal{O}_X} \widetilde{\mathscr{A}}_X^{[0,1)}[1/p]$$

equivariant for the action of $\varphi$ and $\mathscr{G}_{K_\infty}$. In particular this implies that $f$ factors over the admissible locus. □

**Theorem 8.15.** *There is a family $\mathcal{E}^{\mathrm{univ}}$ of semistable $\mathscr{G}_K$-representations on $\mathscr{D}^{\mathrm{ad,adm}}_{\varphi,N,\mu}$ such that $D_{\mathrm{st}}(\mathcal{E}) = (D, \Phi, N, \mathcal{F}^\bullet)$ is the universal family of filtered $(\varphi, N)$-modules on $\mathscr{D}^{\mathrm{ad,adm}}_{\varphi,N,\mu}$. This family is universal in the following sense: Let $X$ be an adic space locally of finite type over $E_\mu$ and let $\mathcal{E}'$ be a family of semistable $\mathscr{G}_K$-representations on $X$ with constant Hodge polygon equal to $\mu$. Then there is a unique morphism $f : X \to \mathscr{D}^{\mathrm{ad,adm}}_{\varphi,N,\mu}$ such that $\mathcal{E}' \cong f^*\mathcal{E}$ as families of $\mathscr{G}_K$-representations.*

*Proof.* The existence of the family $\mathcal{E}$ follows by applying Proposition 8.13 to the family $(\mathfrak{M}, \Phi)$ of $\varphi$-modules of finite height over $\mathscr{A}^{[0,1)}$ on

$$Y = (P_{K_0,d} \times \mathrm{Flag}_{K,d,\mu})^{\mathrm{ad,adm}}.$$

As the construction is obviously functorial, this vector bundle is equivariant for the action of the group $(\mathrm{Res}_{K_0/\mathbb{Q}_p} \mathrm{GL}_{d,K_0})_{E_\mu}$ and hence defines the desired family of semistable $\mathscr{G}_K$-representations on $\mathscr{D}^{\mathrm{ad,adm}}_{\varphi,N,\mu}$. Further the isomorphism $D_{\mathrm{st}}(\mathcal{E}) \cong (D, \Phi, N, \mathcal{F}^\bullet)$ on $Y$ is by construction equivariant under the action of $(\mathrm{Res}_{K_0/\mathbb{Q}_p} \mathrm{GL}_{d,K_0})_{E_\mu}$ and hence descends to $\mathscr{D}^{\mathrm{ad,adm}}_{\varphi,N,\mu}$.

Now let $X$ be as above. The $K$-filtered $(\varphi, N)$-module $D_{\mathrm{st}}(\mathcal{E}')$ defines a morphism $f : X \to \mathscr{D}^{\mathrm{ad}}_{\varphi,N,\mu}$. This map factors over $\mathscr{D}^{\mathrm{ad,adm}}_{\varphi,N,\mu}$ by Proposition 8.14 as factoring over an open subspace may be check on the reduced space underlying $X$. Further we have isomorphisms $D_{\mathrm{st}}(\mathcal{E}') \cong f^* D_{\mathrm{st}}(\mathcal{E}) \cong D_{\mathrm{st}}(f^*\mathcal{E})$. Now the claim follows from Lemma 8.12. □

**Corollary 8.16.** *There is a family $\mathcal{E}$ of crystalline $\mathscr{G}_K$-representations on $\mathscr{D}^{\mathrm{ad,adm}}_{\varphi,\mu}$ such that $D_{\mathrm{cris}}(\mathcal{E}) = (D, \Phi, \mathcal{F}^\bullet)$ is the universal family of filtered $\varphi$-modules on $\mathscr{D}^{\mathrm{ad,adm}}_{\varphi,\mu}$. This family is universal in the following sense: Let $X$ be an adic space locally of finite type over $E_\mu$ and let $\mathcal{E}'$ be a family of crystalline $\mathscr{G}_K$-representations on $X$ with constant Hodge polygon $\mu$. Then there is a unique morphism $f : X \to \mathscr{D}^{\mathrm{ad,adm}}_{\varphi,\mu}$ such that $\mathcal{E}' \cong f^*\mathcal{E}$ as families of $\mathscr{G}_K$-representations.*

*Proof.* This is a direct consequence of the discussion of the semistable case in Theorem 8.15 and Lemma 8.10. □

Let us compare this result to the construction of the universal semistable deformation rings as in [Kisin 2008]. Fix a continuous representation

$$\bar{\rho} : \mathcal{G}_K \to \mathrm{GL}_n(\mathbb{F})$$

with $\mathbb{F}$ a finite extension of $\mathbb{Q}_p$ and write $R_{\bar{\rho}}$ for the universal framed deformation ring of $\bar{\rho}$. Further we write $R_{\bar{\rho}}^{\mathbf{k}}$ for the quotient of $R_{\bar{\rho}}$ constructed in [Kisin 2008, Theorem 2.5.5]. Moreover let us write $\widetilde{\mathscr{D}}^{\mathrm{ad,adm}}_{\varphi,N,\mu}$ for the stack over $\mathscr{D}^{\mathrm{ad,adm}}_{\varphi,N,\mu}$ parametrizing trivializations of the universal semistable representation constructed in Theorem 8.15, i.e., for the stack that assigns to $f : S \to \mathscr{D}^{\mathrm{ad,adm}}_{\varphi,N,\mu}$ the set of isomorphisms $\mathcal{O}_S^n \cong f^*\mathcal{E}^{\mathrm{univ}}$. Note that $\widetilde{\mathscr{D}}^{\mathrm{ad,adm}}_{\varphi,N,\mu}$ actually is representable by an adic space locally of finite type over $\mathbb{Q}_p$ (resp. over the reflex field of $\mu$). We write $\widetilde{\mathscr{D}}^{\mathrm{ad,adm},+}_{\varphi,N,\mu}$ for the open subspace where the canonical representation

$$\rho^{\mathrm{univ}} : \mathcal{G}_K \to \mathrm{GL}(\Gamma(\widetilde{\mathscr{D}}^{\mathrm{ad,adm}}_{\varphi,N,\mu}, \mathcal{E})) \to \mathrm{GL}_n(\Gamma(\widetilde{\mathscr{D}}^{\mathrm{ad,adm}}_{\varphi,N,\mu}, \mathcal{O}))$$

factors over

$$\mathrm{GL}_n(\Gamma(\widetilde{\mathscr{D}}^{\mathrm{ad,adm}}_{\varphi,N,\mu}, \mathcal{O}^+)) \subset \mathrm{GL}_n(\Gamma(\widetilde{\mathscr{D}}^{\mathrm{ad,adm}}_{\varphi,N,\mu}, \mathcal{O})).$$

Note that this really defines an open subspace as the group $\mathcal{G}_K$ is topologically finitely generated (and hence we only need to check for finitely many elements of $\mathscr{G}_K$ whether the corresponding matrix has bounded entries).

Having fixed $\bar{\rho}$ we can cut out an open subspace $\widetilde{\mathscr{D}}^{\mathrm{ad,adm},+}_{\varphi,N,\mu}(\bar{\rho})$ by demanding that the composition

$$\mathcal{G}_K \to \mathrm{GL}_n(\Gamma(\widetilde{\mathscr{D}}^{\mathrm{ad,adm}}_{\varphi,N,\mu}, \mathcal{O}^+)) \to \mathrm{GL}_n(\Gamma(\widetilde{\mathscr{D}}^{\mathrm{ad,adm}}_{\varphi,N,\mu}, \mathcal{O}^+/\mathcal{O}^{++}))$$

is equal to $\bar{\rho}$. Here $\mathcal{O}^{++} \subset \mathcal{O}^+$ denotes the ideal of topologically nilpotent elements. More precisely, given any affinoid open subset $U = \mathrm{Spa}(A, A^+) \subset \widetilde{\mathscr{D}}^{\mathrm{ad,adm},+}_{\varphi,N,\mu}$ we have a canonical family of $\mathcal{G}_K$-representations on the reduced special fiber $\mathrm{Spec}\, A^+/A^{++}$ of $\mathrm{Spf}\, A$ (where $A^{++} \subset A^+$ is the ideal of topologically nilpotent elements), namely

$$\mathcal{G}_K \to \mathrm{GL}_n(A^+) \to \mathrm{GL}_n(A^+/A^{++}).$$

We let $U(\bar{\rho}) \subset U$ denote the tube over the Zariski closed subset of $\mathrm{Spec}\, A^+/A^{++}$ where this composition is equal to $\bar{\rho}$ (or the base change of $\bar{\rho}$ to $A^+/A^{++}$). This construction is obviously compatible with localization on the generic fiber (i.e., with replacing $\mathrm{Spf}\, A^+$ by an affine open subset of an admissible blow up) and hence the pieces $U(\bar{\rho})$ glue together to give $\widetilde{\mathscr{D}}^{\mathrm{ad,adm},+}_{\varphi,N,\mu}(\bar{\rho})$.

Moreover the restriction of $\rho^{\mathrm{univ}}$ to $\widetilde{\mathscr{D}}^{\mathrm{ad,adm},+}_{\varphi,N,\mu}(\bar{\rho})$ induces by construction a morphism to $(\mathrm{Spf}\, R_{\bar{\rho}})^{\mathrm{ad}}$.

**Proposition 8.17.** *The canonical morphism*

$$\widetilde{\mathscr{D}}^{\mathrm{ad,adm},+}_{\varphi,N,\mu}(\bar{\rho}) \to (\mathrm{Spf}\, R_{\bar{\rho}})^{\mathrm{ad}} \tag{8-5}$$

*induces an isomorphism*

$$\widetilde{\mathscr{D}}^{\mathrm{ad,adm},+}_{\varphi,N,\mu}(\bar{\rho}) \cong (\mathrm{Spf}\, R_{\bar{\rho}}^{\mathbf{k}})^{\mathrm{ad}}$$

*Proof.* It follows from [Kisin 2008, Theorem 2.5.5] and the reducedness of the source that the morphism factors over $(\operatorname{Spf} R^{\mathbf{k}}_{\bar\rho})^{\mathrm{ad}}$ and is a bijection on $L$-valued points. Now [Kisin 2008, Theorem 2.5.5] again and the functorial description of the left-hand side show that the morphism is an isomorphism on $A$-valued point for all finite dimensional $\mathbb{Q}_p$-algebras $A$. As the left-hand side is known to be representable we find that the morphism

$$f : \widetilde{\mathscr{D}}^{\mathrm{ad},\mathrm{adm},+}_{\varphi,N,\mu}(\bar\rho) \to (\operatorname{Spf} R^{\mathbf{k}}_{\bar\rho})^{\mathrm{ad}}$$

is a smooth and bijective map of adic spaces locally of finite type over $\mathbb{Q}_p$. Especially it is étale and hence locally given by the composition of an open embedding with a finite étale morphism. As $f$ is bijective on $L$-valued points, the finite étale morphism has to be of degree 1, i.e., an isomorphism. We deduce that $f$ is an open embedding. We conclude that $f$ is an isomorphism by constructing a continuous section to $f$.

Indeed, Kisin's construction [2008, (2.5)] consists of two steps: first he constructs a quotient $A$ of $R_{\bar\rho}$ where the restriction of the universal $\mathscr{G}_K$ representation to $\mathscr{G}_{K_\infty}$ is defined by a $\varphi$-module $\mathfrak{M}$ over $A_W[\![u]\!]$, that is by a $\mathscr{A}^{[0,1)}_{(\operatorname{Spf} A)^{\mathrm{ad}}}$-module of finite height. Let us write $X = (\operatorname{Spf} A)^{\mathrm{ad}}$ and $\mathcal{E}$ for the restriction of the universal $\mathscr{G}_K$-representation to $X$. Then $R^{\mathbf{k}}_{\bar\rho}$ is the quotient of $A$, defined by the condition that the isomorphism

$$\mathfrak{M} \otimes_{\mathscr{A}^{[0,1)}_X} \widetilde{\mathscr{A}}^{[0,1)}_X \Big[\frac{1}{p}\Big] \cong \mathcal{E} \otimes_{\mathcal{O}_X} \widetilde{\mathscr{A}}^{[0,1)}_X \Big[\frac{1}{p}\Big]$$

extends to a $(\mathscr{G}_K, \varphi, N)$-equivariant isomorphism

$$\mathfrak{M} \otimes_{\mathscr{A}^{[0,1)}_X} (B \otimes_{\mathbb{Q}_p} B^+_{\mathrm{st}}) \cong \mathcal{E} \otimes_{\mathcal{O}_X} B \otimes_{\mathbb{Q}_p} B^+_{\mathrm{st}}$$

for every map $R^{\mathbf{k}}_{\bar\rho} \to B$ to a finite dimensional $\mathbb{Q}_p$-algebra. Using Lemma 7.5 (a) and the matrices of the action of $g \in \mathscr{G}_K$ (resp. of $\varphi$ and $N$) in some chosen basis, we deduce that hence the induced isomorphism

$$\mathfrak{M} \otimes_{\mathscr{A}^{[0,1)}_X} (\mathcal{O}_X \widehat{\otimes} B^+_{\mathrm{st}}) \cong \mathcal{E} \otimes_{\mathcal{O}_X} \mathcal{O}_X \widehat{\otimes} B^+_{\mathrm{st}}$$

is equivariant for the actions of $(\mathscr{G}_K, \varphi, N)$. In particular the family of Galois representations on $(\operatorname{Spf} R^{\mathbf{k}}_{\bar\rho})^{\mathrm{ad}}$ is semistable according to our definition.

Hence we obtain a canonical morphism

$$(\operatorname{Spf} R^{\mathbf{k}}_{\bar\rho})^{\mathrm{ad}} \to \mathscr{D}^{\mathrm{ad},\mathrm{adm}}_{\varphi,N,\mu}.$$

As $\mathcal{E}$ comes with a trivialization of an $\mathcal{G}_K$-stable $\mathcal{O}^+$-lattice inside $\mathcal{E}$ this morphism canonically lifts to $\widetilde{\mathscr{D}}^{\mathrm{ad},\mathrm{adm},+}_{\varphi,N,\mu}(\bar\rho)$ and defines a morphism that is set-theoretically a section to $f$. As $f$ already is known to be an open embedding it is enough to conclude. □

**Remark 8.18.** We note that Kisin's description of the semistable deformation rings is a priori quite different: a family of Galois representations over some affinoid algebra $A$ is crystalline (resp. semistable) in Kisin's sense if it is crystalline (resp. semistable) after the base change to each quotient of $A$ that is finite-dimensional as a $\mathbb{Q}_p$-vector space. On the other hand we have aimed at giving a definition of a family of crystalline representations in the spirit of Fontaine (though we did not do this using filtered

$\varphi$-modules, but rather $\varphi$-modules on the open unit disc). As it is not so obvious how these definitions directly relate to each other we construct the morphism (8-5) in a slightly complicated manner.

Note that our construction has the advantage that we no longer need to fix a framing of an integral structure inside the Galois representation. After this paper was written, Wang-Erickson [2018] extended the results of Kisin in a direct way to families that do not longer fix a framing. For such families a similar comparison with our construction should hold true. However, it seems that one cannot recover the main result of [Wang-Erickson 2018] from our construction that takes place purely in the generic fiber.

We finally comment on the relation of our construction with the work of Berger and Colmez [2008]. They studied families of $p$-adic representations parametrized by $p$-adic Banach algebras. They proved for example that in such a family the locus of point-wise crystalline (resp. semistable) representations of fixed Hodge–Tate weight is a closed subspace, and there exist a family of filtered $\varphi$-modules (resp. filtered $(\varphi, N)$-modules) that specializes to the filtered $(\varphi, N)$-modules at each point. We deduce from the comparison with Kisin's construction that our families have the same property.

**Corollary 8.19.** *Let $X$ be a reduced adic space locally of finite type over $\mathbb{Q}_p$ and let $\mathcal{E}$ be a family of $\mathscr{G}_K$-representations on $X$. We assume that (fpqc-locally on $X$) there exists a $\mathscr{G}_K$-stable $\mathcal{O}_X^+$-lattice in $\mathcal{E}$. Then $\mathcal{E}$ is a semistable family if and only if $\mathcal{E}\otimes\kappa(x)$ is semistable for all $x\in X$.*

*Proof.* The subspace $(\operatorname{Spf} R_{\bar\rho}^{\mathbf{k}})^{\mathrm{ad}} \subset (\operatorname{Spf} R_{\bar\rho})^{\mathrm{ad}}$ is the Zariski-closure of all classical points at which the universal Galois representation on $(\operatorname{Spf} R_{\bar\rho})^{\mathrm{ad}}$ is semistable with Hodge–Tate weight $\mathbf{k}$. The result hence follows from the fact that by assumption we may (locally in the fpqc-topology) construct a morphism $X \to (\operatorname{Spf} R_{\bar\rho})^{\mathrm{ad}}$ such that the pullback of the universal representation on $(\operatorname{Spf} R_{\bar\rho})^{\mathrm{ad}}$ agrees with the $\mathscr{G}_K$-stable $\mathcal{O}_X^+$ lattice in $\mathcal{E}$. □

We remark that the existence of an integral lattice is always assumed in [Berger and Colmez 2008]. In fact we do not know whether it automatically exists or whether this is a true condition. As our definition (in particular the definition of the completed sheaves of period rings) differs, the relation of our construction with theirs is less clear in the nonreduced case. However, the universal case is reduced.

## 9. The morphism to the adjoint quotient

As in [Hellmann 2011, §4] we consider the adjoint quotient $A/\mathfrak{S}_d$, where $A\subset \mathrm{GL}_{d,\mathbb{Q}_p}$ is the diagonal torus and $\mathfrak{S}_d$ is the finite Weyl group of $\mathrm{GL}_d$. Under the morphism $c : A \to \mathbb{A}^{d-1}_{\mathbb{Q}_p} \times_{\mathbb{Q}_p} \mathbb{G}_{m,\mathbb{Q}_p}$ which maps an element $g$ of $A$ to the coefficients $c_1,\dots,c_d$ of its characteristic polynomial $\chi_g = X^d + c_1X^{d-1} + \cdots + c_d$, the adjoint quotient $A/\mathfrak{S}_d$ is isomorphic to $\mathbb{A}^{d-1}_{\mathbb{Q}_p} \times_{\mathbb{Q}_p} \mathbb{G}_{m,\mathbb{Q}_p} = \operatorname{Spec}\mathbb{Q}_p[c_1,\dots,c_d,c_d^{-1}]$. Recall from [Hellmann 2011, §4] that there is a morphism

$$\operatorname{Res}_{K_0/\mathbb{Q}_p}\mathrm{GL}_{d,K_0} \to A/\mathfrak{S}_d \tag{9-1}$$

which is invariant under $\varphi$-conjugation on the source. It is defined on $R$-valued points by sending $b\in(\operatorname{Res}_{K_0/\mathbb{Q}_p}\mathrm{GL}_{d,K_0})(R)=\mathrm{GL}_d(R\otimes_{\mathbb{Q}_p}K_0)$ to the characteristic polynomial of $(b\cdot\varphi)^f = b\cdot\varphi(b)\cdots\varphi^{(f-1)}(b)$,

where $f = [K_0 : \mathbb{Q}_p]$. This characteristic polynomial actually has coefficients in $R$, because it is invariant under $\varphi$, as can be seen from the formula $\varphi(b \cdot \varphi)^f = b^{-1} \cdot (b \cdot \varphi)^f \cdot \varphi^f(b) = b^{-1} \cdot (b \cdot \varphi)^f \cdot b$. Since $\mathrm{Res}_{K_0/\mathbb{Q}_p} \mathrm{GL}_{d,K_0}$ acts on itself by $\varphi$-conjugation via $(g, b) \mapsto g^{-1} b\, \varphi(g)$ and $(g^{-1} b\, \varphi(g)\, \varphi)^f = g^{-1} \cdot (b \cdot \varphi)^f \cdot g$ the map (9-1) is invariant under $\varphi$-conjugation.

Let $\mu$ be a cocharacter as in (2-5), let $E_\mu$ be its reflex field, and set $(A/\mathfrak{S}_d)_{E_\mu} := A/\mathfrak{S}_d \times_{\mathbb{Q}_p} E_\mu$. By projecting to $\mathrm{Res}_{K_0/\mathbb{Q}_p} \mathrm{GL}_{d,K_0}$ we may extend $\beta$ to morphisms

$$\begin{array}{ccc} P_{K_0,d} \times_{\mathbb{Q}_p} Q_{K,d,\preceq\mu} & \xrightarrow{\widetilde{\alpha}} & (A/\mathfrak{S}_d)_{E_\mu} \\ \downarrow & & \| \\ \mathscr{H}_{\varphi,N,\preceq\mu} & \xrightarrow{\alpha} & (A/\mathfrak{S}_d)_{E_\mu} \end{array}$$

We further obtain morphisms to $(A/\mathfrak{S}_d)_{E_\mu}$ from the locally closed substacks $\mathscr{H}_{\varphi,\preceq\mu}$, $\mathscr{H}_{\varphi,N,\mu}$, $\mathscr{H}_{\varphi,\mu}$, $\mathscr{D}_{\varphi,N,\mu}$, and $\mathscr{D}_{\varphi,\mu}$, which we likewise denote by $\alpha$. Here we view $\mathscr{D}_{\varphi,N,\mu}$ and $\mathscr{D}_{\varphi,\mu}$ as substacks of $\mathscr{H}_{\varphi,N,\mu}$ via the zero section from Remark 2.8 (3). We also consider the adification of these morphisms.

**Theorem 9.1.** *Let $\mu$ be a cocharacter as in* (2-5), *let $E_\mu$ be its reflex field and let $x \in (A/\mathfrak{S}_d)^{\mathrm{ad}}_{E_\mu}$. Then there exists an open subscheme $X$ of $\widetilde{\alpha}^{-1}(x)$ such that the weakly admissible locus in the fiber over $x$ is given by*

$$\widetilde{\alpha}^{-1}(x)^{\mathrm{wa}} = X^{\mathrm{ad}}.$$

*Proof.* This is similar to the proof of [Hellmann 2011, Theorem 4.1]. Let

$$x = (c_1, \ldots, c_d) \in \kappa(x)^{d-1} \times \kappa(x)^\times$$

and let $v_x$ denote the (multiplicative) valuation on $\kappa(x)$. First note that

$$c_d = \det_{\kappa(x) \otimes_{\mathbb{Q}_p} K_0}(b \cdot \varphi)^f = \det_{\kappa(x)}((b \cdot \varphi)^f)^{1/f}$$

and hence $\widetilde{\alpha}^{-1}(x)^{\mathrm{wa}} = \varnothing$ unless

$$v_x(c_d)^{-1/f} \cdot v_x(p)^{\frac{1}{ef} \sum_{\psi,j} \mu_{\psi,j}} = \lambda(\underline{D}) = 1.$$

In the following we will assume that this condition is satisfied. We now revert to the notation of the proof of Theorem 5.6. In particular we consider the projective $P_{K_0,d}$-schemes $Z_i$, the global sections $f_i \in \Gamma(Z_i, \mathcal{O}_{Z_i})$, the functions $h_i$, the closed subsets

$$Y_{i,m} = \{y \in Z_i^{\mathrm{ad}} \times Q^{\mathrm{ad}}_{K,d,\preceq\mu} \mid h_i(y) \geq m\}$$

and the proper projections $\mathrm{pr}_{i,m} : Y_{i,m} \to P_{K_0,d} \times_{\mathbb{Q}_p} Q_{K,d,\preceq\mu}$. This time

$$S_{i,m} = \{y = (g_y, N_y, U_y, \mathfrak{q}_y) \in Y_{i,m} \times_{(P_{K_0,d} \times Q_{K,d,\preceq\mu})} \widetilde{\alpha}^{-1}(x) \mid v_y(f_i(g_y, U_y)) > v_y(p)^{f^2 m}\}$$

is a union of connected components of $Y_{i,m} \times_{(P_{K_0,d} \times Q_{K,d,\preceq\mu})} \widetilde{\alpha}^{-1}(x)$; hence a closed subscheme and not just a closed adic subspace. This can be seen as follows: Let $\lambda_1, \ldots, \lambda_d$ denote the zeros of the polynomial

$$X^d + c_1 X^{d-1} + \cdots + c_{d-1} X + c_d.$$

Then every possible value of the $f_i$ is a product of some of the $\lambda_i$ and hence $f_i$ can take only finitely many values. As in the proof of Theorem 5.6

$$\widetilde{\alpha}^{-1}(x)^{\mathrm{wa}} = \widetilde{\alpha}^{-1}(x) \smallsetminus \bigcup_{i,m} \mathrm{pr}_{i,m}(S_{i,m}),$$

where the union runs over $1 \le i \le d-1$ and $m \in \mathbb{Z}$. So $\widetilde{\alpha}^{-1}(x)^{\mathrm{wa}}$ is an open subscheme of $\widetilde{\alpha}^{-1}(x)$. □

**Corollary 9.2.** *Let $x \in (A/\mathfrak{S}_d)^{\mathrm{ad}}_E$ and consider the 2-fiber product*

$$\begin{array}{ccc} \alpha^{-1}(x)^{\mathrm{wa}} & \longrightarrow & \mathscr{H}^{\mathrm{ad,wa}}_{\varphi,N,\preceq\mu} \\ \downarrow & & \downarrow \alpha \\ x & \longrightarrow & (A/W)^{\mathrm{ad}}_E \end{array}$$

*Then there exists an Artin stack in schemes $\mathfrak{A}$ over the field $\kappa(x)$ which is an open substack of $\alpha^{-1}(x)$, such that $\alpha^{-1}(x)^{\mathrm{wa}} = \mathfrak{A}^{\mathrm{ad}}$. The same is true for $\mathscr{H}_{\varphi,\preceq\mu}$, $\mathscr{H}_{\varphi,N,\mu}$, $\mathscr{H}_{\varphi,\mu}$, $\mathscr{D}_{\varphi,N,\mu}$, and $\mathscr{D}_{\varphi,\mu}$.*

*Proof.* This is an immediate consequence of Theorem 9.1 and the proof of Corollary 5.7. □

We also determine the image of the weakly admissible locus in the adjoint quotient.

**Theorem 9.3.** *The image of $\mathscr{H}^{\mathrm{ad,wa}}_{\varphi,N,\preceq\mu}$ (and $\mathscr{H}^{\mathrm{ad,wa}}_{\varphi,\preceq\mu}$, $\mathscr{H}^{\mathrm{ad,wa}}_{\varphi,N,\mu}$, $\mathscr{H}^{\mathrm{ad,wa}}_{\varphi,\mu}$, $\mathscr{D}^{\mathrm{ad,wa}}_{\varphi,N,\mu}$, and $\mathscr{D}^{\mathrm{ad,wa}}_{\varphi,\mu}$) under the morphism(s) $\alpha$ is equal to the affinoid subdomain*

$$\left\{ c=(c_1,\dots,c_d) \in (A/\mathfrak{S}_d)^{\mathrm{ad}}_{E_\mu} \;\middle|\; v_c(c_i) \le v_c(p)^{\frac{1}{e}\sum_\psi(\mu_{\psi,d}+\cdots+\mu_{\psi,d+1-i})} \text{ with equality for } i=d \right\}, \quad (9\text{-}2)$$

*where $v_c$ is the (multiplicative) valuation of the adic point $c$ with $v_c(p) < 1$.*

**Remark 9.4.** (1) The subset described in (9-2) is really an affinoid subdomain. Indeed the adjoint quotient $(A/\mathfrak{S}_d)^{\mathrm{ad}}_{E_\mu}$ is (admissibly) covered by the (admissible) open affinoid rigid spaces (or adic spaces)

$$X_M = \left\{ c = (c_1,\dots,c_d) \in (A/\mathfrak{S}_d)^{\mathrm{ad}}_{E_\mu} \mid v_c(c_i) \le p^M, \text{ for all } i \text{ and } v_c(v_d) \ge -p^M \right\}$$

and the subspace (9-2) is easily seen to be a Laurent subdomain of each of these $X_M$ for $M \gg 0$.

(2) The morphisms $\alpha$ forget the Hodge–Pink lattice $\mathfrak{q}$ (or the $K$-filtration $\mathcal{F}^\bullet$) and in general their fibers contain infinitely many weakly admissible points.

(3) Like in [Hellmann 2011, Proposition 5.2] the affinoid subdomain of Theorem 9.3 can be described as the *closed Newton stratum* of the coweight $\left(-\frac{1}{e}\sum_\psi \mu_{\psi,d} \ge \cdots \ge -\frac{1}{e}\sum_\psi \mu_{\psi,1}\right)$ of $A$. By this we mean that the $\overline{\mathbb{Q}}_p$-valued points (i.e., the rigid analytic points) of (9-2) coincide with the points of the corresponding Newton stratum in the sense of [Kottwitz 2006]. In [Hellmann 2011] the claim is made for all points of the corresponding Berkovich space. In the set up of adic spaces we cannot rely on Kottwitz's definition of a Newton stratum for all points of the adic space, as the valuations are not necessarily rank one valuations, i.e., the value group is not necessarily a subgroup of the real numbers. Especially the Newton strata do not cover the adic space $(A/\mathfrak{S}_d)^{\mathrm{ad}}$.

(4) For $\mathscr{D}^{\mathrm{ad,wa}}_{\varphi,\mu}$ the description of the image in our Theorem 9.3 has previously been obtained by Fontaine and Rapoport [2005, Théorèm 1] and Breuil and Schneider [2007, Proposition 3.2] on the level of $L$-valued points where in [Fontaine and Rapoport 2005] $L$ is a complete discretely valued extension of $E_\mu$ with algebraically closed residue field. In [Breuil and Schneider 2007] $L$ is a finite extension of $E_\mu$ and in addition all Hodge–Tate weights are assumed to be pairwise different. Moreover, our affinoid subdomain (9-2) equals $\mathfrak{S}_d\backslash\mathbb{T}'_\xi$ from [Breuil and Schneider 2007, Corollary 2.5], where $\xi$ is associated with the cocharacter $\tilde{\xi} := \big(-\mu-(0,1,\dots,d-1)\big)_{\mathrm{dom}} \in X_*(\widetilde{T})$. Actually, both [Fontaine and Rapoport 2005; Breuil and Schneider 2007] even prove that over an $L$-valued point $c$ in the image there is an $L$-valued point in $\widetilde{\alpha}^{-1}(c)^{\mathrm{wa}}$. This also follows from our Theorem 9.1, which shows that $\widetilde{\alpha}^{-1}(c)^{\mathrm{wa}}$ is Zariski-open in a scheme covered by affine spaces, see (3-2), because the $L$-valued points (for any infinite field $L$) lie dense in such schemes. In this way our theorem provides a new proof for [Fontaine and Rapoport 2005, Théorèm 1] and generalizes [Breuil and Schneider 2007, Proposition 3.2]; see Section 10 for more details.

Before we prove the theorem we note the following:

**Lemma 9.5.** *Set $l_i := \frac{1}{ef}\sum_\psi(\mu_{\psi,d}+\dots+\mu_{\psi,d+1-i})$. Then $l_i$ equals the number $l_i$ defined in* [*Hellmann 2011, Formula* (5.2) *on p.* 988]. *If $\underline{D}=(D,\Phi,N,\mathfrak{q})$ is a $(\varphi,N)$-module with Hodge–Pink lattice over a field $L\supset E_\mu$ whose Hodge polygon is bounded by $\mu$ and if $\underline{D}'=\big(D',\Phi|_{\varphi^*D'},N|_{D'},\mathfrak{q}\cap D'\otimes_{L\otimes K_0}\mathbb{B}_L\big)\subset\underline{D}$ for a free $L\otimes_{\mathbb{Q}_p}K_0$-submodule $D'\subset D$ of rank $i$ which is stable under $\Phi$ and $N$, then $t_H(\underline{D}')\geq l_i$.*

*Proof.* The number $l_i$ in [Hellmann 2011, (5.2)] was defined as follows. Write $\{\mu_{\psi,1},\dots,\mu_{\psi,d}\}=\{x_{\psi,1},\dots,x_{\psi,r}\}$ with $x_{\psi,j}>x_{\psi,j+1}$. Let $n_{\psi,j}:=\max\{k:\mu_{\psi,k}\geq x_{\psi,j}\}$. In particular $n_{\psi,r}=d$ and $\mu_{\psi,n_{\psi,j}}\geq x_{\psi,j}$. For $0\leq i\leq d$ let $m_{\psi,j}(i):=\max\{0,n_{\psi,j}+i-d\}$. So $m_{\psi,j}(0)=0$ for all $j$ and $m_{\psi,r}(i)=i$. It follows that $n_{\psi,j}\geq d-i$ if and only if $\mu_{\psi,d-i}\geq x_{\psi,j}$. Now $l_i$ was defined in [Hellmann 2011, (5.2)] as

$$l_i=\frac{1}{ef}\sum_\psi\Big(\sum_{j=1}^{r-1}(x_{\psi,j}-x_{\psi,j+1})m_{\psi,j}(i)+x_{\psi,r}m_{\psi,r}(i)\Big).$$

We compute

$$l_{i+1}-l_i=\frac{1}{ef}\sum_\psi\Big(\sum_{j=1}^{r-1}(x_{\psi,j}-x_{\psi,j+1})\big(m_{\psi,j}(i+1)-m_{\psi,j}(i)\big)+x_{\psi,r}\Big).$$

The difference $m_{\psi,j}(i+1)-m_{\psi,j}(i)$ is 1 if $n_{\psi,j}+i-d\geq 0$, that is if $x_{\psi,j}\leq\mu_{\psi,d-i}$. Otherwise the difference $m_{\psi,j}(i+1)-m_{\psi,j}(i)$ is 0. Therefore $l_{i+1}-l_i=\frac{1}{ef}\sum_\psi\mu_{\psi,d-i}$ and $l_0=0$ implies that $l_i=\frac{1}{ef}\sum_\psi(\mu_{\psi,d}+\dots+\mu_{\psi,d+1-i})$.

To prove the second assertion let $s\in\operatorname{Spec}L\otimes_{E_\mu}\widetilde{K}$ be a point and let $\mu'=\mu_{\underline{D}}(s)$ be the Hodge polygon of $\underline{D}$ at $s$. Then $\mu_{\psi,d}+\dots+\mu_{\psi,d+1-i}\leq\mu'_{\psi,d}+\dots+\mu'_{\psi,d+1-i}$ for all $\psi$ and all $i$ by Proposition 2.13(b). We let $\mathfrak{p}_\psi$ be the $\psi$-component of $s^*\mathfrak{p}:=s^*D\otimes_{\kappa(s)\otimes K_0}\mathbb{B}^+_{\kappa(s)}$ and $\mathfrak{p}'_\psi$ be the $\psi$-component of $s^*\mathfrak{p}':=s^*D'\otimes_{\kappa(s)\otimes K_0}\mathbb{B}^+_{\kappa(s)}$. By definition of the Hodge polygon, see Construction 2.10, we can choose a

$\kappa(s)[\![t]\!]$-basis $(v_{\psi,1},\dots,v_{\psi,d})$ of $\mathfrak{p}_\psi$ such that $(t^{-\mu'_{\psi,1}}v_{\psi,1},\dots,t^{-\mu'_{\psi,d}}v_{\psi,d})$ is a $\kappa(s)[\![t]\!]$-basis of the $\psi$-component $\mathfrak{q}_\psi$ of $s^*\mathfrak{q}$. Since

$$\dim_{\kappa(s)(\!(t)\!)}\Big(\mathfrak{p}'_\psi\Big[\frac{1}{t}\Big]\cap\langle v_{\psi,1},\dots,v_{\psi,n}\rangle_{\kappa(s)(\!(t)\!)}\Big)\geq n+i-d$$

for all $n$, we can find a $\kappa(s)[\![t]\!]$-basis $(v'_{\psi,1},\dots,v'_{\psi,i})$ of $\mathfrak{p}'_\psi$ with $v'_{\psi,j}\in\langle v_{\psi,1},\dots,v_{\psi,d+j-i}\rangle_{\kappa(s)[\![t]\!]}$. Namely, for each $j$ we let $\bar{v}'_j$ be an element of

$$\big(\mathfrak{p}'_\psi\cap\langle v_{\psi,1},\dots,v_{\psi,d+j-i}\rangle_{\kappa(s)[\![t]\!]}\big)\big/\langle v'_{\psi,1},\dots,v'_{\psi,j-1}\rangle_{\kappa(s)[\![t]\!]}$$

which generates a nonzero saturated $\kappa(s)[\![t]\!]$-submodule, and we let $v'_{\psi,j}\in\mathfrak{p}'_\psi\cap\langle v_{\psi,1},\dots,v_{\psi,d+j-i}\rangle_{\kappa(s)[\![t]\!]}$ be a lift of $\bar{v}'_{\psi,j}$. Then $(v'_{\psi,1},\dots,v'_{\psi,j})$ is linearly independent over $\kappa(s)(\!(t)\!)$ and generates a saturated $\kappa(s)[\![t]\!]$-submodule of $\mathfrak{p}'_\psi$. Using this basis we see that $t^{-\mu'_{\psi,d+j-i}}\cdot v'_{\psi,j}\in\mathfrak{q}_\psi\cap\mathfrak{p}'_\psi[1/t]$. This implies that $t_H(\underline{D}')\geq\frac{1}{ef}\sum_\psi\mu'_{\psi,d}+\cdots+\mu'_{\psi,d+1-i}\geq l_i$. □

*Proof of Theorem 9.3.* We consider the embedding of $\mathscr{D}_{\varphi,\mu}$ into $\mathscr{H}_{\varphi,\mu}$ via the zero section. Under this section $\mathscr{D}^{\mathrm{ad,wa}}_{\varphi,\mu}$ is contained in $\mathscr{H}^{\mathrm{ad,wa}}_{\varphi,N,\preceq\mu}$, $\mathscr{H}^{\mathrm{ad,wa}}_{\varphi,\preceq\mu}$, $\mathscr{H}^{\mathrm{ad,wa}}_{\varphi,N,\mu}$, $\mathscr{H}^{\mathrm{ad,wa}}_{\varphi,\mu}$, and $\mathscr{D}^{\mathrm{ad,wa}}_{\varphi,N,\mu}$ by Lemma 5.3 or Remark 5.8. Conversely they are all contained in $\mathscr{H}^{\mathrm{ad,wa}}_{\varphi,N,\preceq\mu}$. Moreover, these inclusions are compatible with the morphisms $\alpha$ to $(A/\mathfrak{S}_d)_{E_\mu}$.

We first claim that the affinoid subdomain is contained in the image of the weakly admissible locus for all these stacks. By the above it suffices to prove the claim for $\mathscr{D}^{\mathrm{ad,wa}}_{\varphi,\mu}$. In this case the claim follows from [Hellmann 2011, Theorem 5.5 and Proposition 5.2] using Lemma 9.5. Note that in [loc. cit.] only Berkovich's analytic points are treated, but the given argument works verbatim also for adic points.

Conversely let $c=(c_1,\dots,c_d)$ be an $L$-valued point of $(A/\mathfrak{S}_d)^{\mathrm{ad}}_{E_\mu}$ which lies in the image of the weakly admissible locus of one of these stacks. By the above it lies in the image of $\mathscr{H}^{\mathrm{ad,wa}}_{\varphi,N,\preceq\mu}$. So let $\underline{D}\in\mathscr{H}^{\mathrm{ad,wa}}_{\varphi,N,\preceq\mu}(L')$ for a field extension $L'/L$, such that $\underline{D}$ maps to $c$. By extending the field $L'$ further we may assume that $K_0\subset L'$ and that $X^d+c_1X^{d-1}+\cdots+c_d=\prod_{j=1}^d(X-\lambda_j)$ splits into linear factors with $\lambda_j\in L'$. We claim that $v_{L'}(\prod_{j\in I}\lambda_j)\leq v_{L'}(p)^{fl_i}$ for all subsets $I\subset\{1,\dots,d\}$ of cardinality $i$. By Lemma 9.5 this implies that $c$ lies in our affinoid subdomain.

To prove the claim we use Remark 2.4. Then $X^d+c_1X^{d-1}+\cdots+c_d$ is the characteristic polynomial of the $L'$-endomorphism $(\Phi^f)_0$ of $D_0$ and $t_N(\underline{D})=v_{L'}(\det_{L'}(\Phi^f)_0)^{1/f}$. We write $(\Phi^f)_0$ in Jordan canonical form and observe that $N_0$ maps the generalized eigenspace of $(\Phi^f)_0$ with eigenvalue $\lambda_j$ into the one with eigenvalue $p^{-f}\lambda_j$. If $I\subset\{1,\dots,d\}$ is a subset with cardinality $i$ this allows us to find an $i$-dimensional $L'$-subspace $D'_0\subset D_0$ which is stable under $(\Phi^f)_0$ and $N_0$ such that the eigenvalues of $(\Phi^f)_0$ on $D'_0$ are of the form $(p^{-n_j}\lambda_j : j\in I)$ for suitable $n_j\in\mathbb{Z}_{\geq 0}$. We let $D'\subset D$ be the $(\varphi,N)$-submodule corresponding to $D'_0\subset D_0$ under Remark 2.4. Then

$$v_{L'}\Big(\prod_{j\in I}\lambda_j\Big)\leq v_{L'}\Big(\prod_{j\in I}p^{-n_j}\lambda_j\Big)=v_{L'}\big(\det_{L'}(\Phi^f)_0|_{D'_0}\big)=t_N(\underline{D}')^f\leq v_{L'}(p)^{f\,t_H(\underline{D}')}\leq v_{L'}(p)^{fl_i}$$

by the weak admissibility of $\underline{D}$ and by Lemma 9.5. This proves the theorem. □

## 10. Applications

Let us mention two conjectural applications of our constructions to the $p$-adic local Langlands program.

***Breuil's conjecture on the locally analytic socle.*** Breuil [2015; 2016] formulated a conjecture on the locally analytic principal series representations that embed into the $\rho$-isotypical part of completed cohomology (or some $p$-adically completed space of automorphic forms) for some fixed global Galois representation $\rho$ which is associated to an automorphic representation. The automorphic representation to which $\rho$ is associated defines a locally algebraic representation inside completed cohomology, i.e., a representation that appears in the conjecture of Breuil and Schneider; see below. The conjectured existence of more locally analytic principal series representations is the representation-theoretic formulation of the existence of *companion points* on eigenvarieties, i.e., the existence of (overconvergent) $p$-adic automorphic forms (of finite slope) such that the associated Galois representation is in fact automorphic.

These additional locally analytic representations that should conjecturally embed into completed cohomology are described by combinatorial data: the relative position of the de Rham filtration and a flag of $\varphi$-stable subspaces inside $D_{\mathrm{st}}(\rho)$, i.e., they are described completely by local data. In fact one can formulate a conjecture for all (potentially) semistable local Galois representations (not just the restrictions of global Galois representations) by replacing the completed cohomology by the candidate for the $p$-adic local Langlands correspondence as in [Caraiani et al. 2016].

In [Breuil et al. 2019], Breuil, Schraen and the second author established a link between the existence of these locally analytic principal series representations and the degenerations of certain structures from $p$-adic Hodge theory (and the theory of $(\varphi, \Gamma)$-modules) in rigid analytic families. The degenerations predicted by Breuil's conjecture can be constructed using precisely the universal families of semistable representations defined in the present article.

***The Breuil–Schneider conjecture.*** This second application is rather a speculation than a true application. As mentioned in the introduction the $p$-adic local Langlands program wants to relate on the one hand certain continuous representations of $\mathscr{G}_K$ on $n$-dimensional $L$-vector spaces for another $p$-adic field $L$, and on the other hand topologically irreducible admissible representations of $\mathrm{GL}_n(K)$ on finite-dimensional $L$-Banach spaces. We want to explain in which sense both kinds of representations vary in families.

On the side of $\mathrm{GL}_n(K)$-representations, when all Hodge–Tate weights are pairwise different, a Banach–Hecke algebra $\mathcal{B}$ which is the completion of the usual spherical Hecke algebra for a certain norm was constructed in [Breuil and Schneider 2007; Schneider and Teitelbaum 2006]. This Banach–Hecke algebra is an affinoid algebra over the Galois closure $\widetilde{K}$ of $K/\mathbb{Q}_p$, whose associated affinoid space $\operatorname{Spa}\mathcal{B}$ is contained in a split $n$-dimensional torus $A$. Moreover, the algebra $\mathcal{B}$ acts on a universal infinite-dimensional locally algebraic Banach representation of $\mathrm{GL}_n(K)$. Breuil and Schneider also conjectured that the specialization of the universal Banach representation at any $L$-valued point of $\operatorname{Spa}\mathcal{B}$ admits an (in general many) invariant norm(s) and proved this in some cases. Further cases were established by Sorensen [2013] and more recently many new cases were proved by Caraiani, Emerton, Gee, Geraghty,

Paškūnas and Shin [Caraiani et al. 2016]. One might hope that the completions with respect to these norms produce the searched for irreducible admissible finite-dimensional $L$-Banach representations.

If on the Galois side one restricts to semistable or crystalline representations of $\mathscr{G}_K$ then we provide in this article the moduli spaces $\mathscr{D}^{\mathrm{ad,adm}}_{\varphi,N,\mu}$ for those. Sending a semistable $\mathscr{G}_K$-representation to the characteristic polynomial of its associated Frobenius defines a morphism $\alpha$ from $\mathscr{D}^{\mathrm{ad,adm}}_{\varphi,N,\mu}$ to the adjoint quotient $(A/\mathfrak{S}_n)^{\mathrm{ad}}$ which contains (an image of) the affinoid domain $\operatorname{Spa}\mathcal{B}$ of Breuil and Schneider. In Section 9 we proved that the fibers of this morphism $\alpha$ are Artin stacks in schemes (Corollary 9.2) and we determined the image of $\alpha$. If all Hodge–Tate weights are pairwise different, Breuil and Schneider [2007, Proposition 3.2] proved that the image equals $\operatorname{Spa}\mathcal{B}$. Our Theorem 9.3 generalizes this to arbitrary Hodge–Tate weights. So one may now ask whether there is a relation between the fiber of the morphism $\alpha$ over an $L$-valued point of $\operatorname{Spa}\mathcal{B}$ and the set of invariant norms on the specialization of the universal Banach representation at this point.

The reader should note that by the condition of [Hellmann 2013] that the Hodge–Tate weights lie in $\{0, 1\}$ together with the condition of [Breuil and Schneider 2007; Schneider and Teitelbaum 2006] that they are pairwise different, one was limited to $\mathrm{GL}_2$ for which the $p$-adic local Langlands program is established when $K = \mathbb{Q}_p$; see [Colmez 2010; Paškūnas 2013; Colmez et al. 2014]. So for the application to $\mathrm{GL}_n$ when $n > 2$ our generalization in the present article is essential.

## Acknowledgements

Hellmann acknowledges support of the DFG (German Research Foundation) in form of SFB/TR 45 "Periods, Moduli Spaces and Arithmetic of Algebraic Varieties" and in form of a Forschungsstipendium He 6753/1-1. Both authors were also supported by SFB 878 "Groups, Geometry & Actions" of the DFG and Germany's Excellence Strategy EXC 2044–390685587 "Mathematics Münster: Dynamics-Geometry-Structure". We would further like to thank A. Mézard, M. Rapoport, T. Richarz and P. Scholze for helpful discussions.

## References

[Altman and Kleiman 1980] A. B. Altman and S. L. Kleiman, "Compactifying the Picard scheme", *Adv. in Math.* **35**:1 (1980), 50–112. MR Zbl

[Beauville and Laszlo 1994] A. Beauville and Y. Laszlo, "Conformal blocks and generalized theta functions", *Comm. Math. Phys.* **164**:2 (1994), 385–419. MR Zbl

[Beilinson and Drinfeld 2005] A. Beilinson and V. Drinfeld, "Quantization of Hitchin's integrable system and Hecke eigensheaves", preprint, 2005, available at http://math.uchicago.edu/~drinfeld/langlands/hitchin/BD-hitchin.pdf.

[Berger 2002] L. Berger, "Représentations $p$-adiques et équations différentielles", *Invent. Math.* **148**:2 (2002), 219–284. MR Zbl

[Berger and Colmez 2008] L. Berger and P. Colmez, "Familles de représentations de de Rham et monodromie $p$-adique", pp. 303–337 in *Représentations $p$-adiques de groupes $p$-adiques*, vol. I: Représentations galoisiennes et $(\varphi, \Gamma)$-modules, Astérisque **319**, Société Mathématique de France, Paris, 2008. MR Zbl

[Bosch and Lütkebohmert 1993] S. Bosch and W. Lütkebohmert, "Formal and rigid geometry, II: Flattening techniques", *Math. Ann.* **296**:3 (1993), 403–429. MR Zbl

[Bosch et al. 1984] S. Bosch, U. Güntzer, and R. Remmert, *Non-Archimedean analysis: a systematic approach to rigid analytic geometry*, Grundl. Math. Wissen. **261**, Springer, 1984. MR Zbl

[Bourbaki 1961] N. Bourbaki, *Algèbre commutative: Chapitre 3 et Chapitre 4*, Actualités Scientifiques et Industrielles **1293**, Hermann, Paris, 1961. MR Zbl

[Bourbaki 1970] N. Bourbaki, *Algèbre: Chapitres 1 à 3*, Hermann, Paris, 1970. MR Zbl

[Breuil 2015] C. Breuil, "Vers le socle localement analytique pour $\mathrm{GL}_n$ II", *Math. Ann.* **361**:3-4 (2015), 741–785. MR Zbl

[Breuil 2016] C. Breuil, "Socle localement analytique I", *Ann. Inst. Fourier* (*Grenoble*) **66**:2 (2016), 633–685. MR Zbl

[Breuil and Schneider 2007] C. Breuil and P. Schneider, "First steps towards $p$-adic Langlands functoriality", *J. Reine Angew. Math.* **610** (2007), 149–180. MR Zbl

[Breuil et al. 2019] C. Breuil, E. Hellmann, and B. Schraen, "A local model for the trianguline variety and applications", *Publ. Math. Inst. Hautes Études Sci.* **130** (2019), 299–412. MR

[Caraiani et al. 2016] A. Caraiani, M. Emerton, T. Gee, D. Geraghty, V. Paškūnas, and S. W. Shin, "Patching and the $p$-adic local Langlands correspondence", *Camb. J. Math.* **4**:2 (2016), 197–287. MR Zbl

[Colmez 2010] P. Colmez, "Représentations de $\mathrm{GL}_2(\mathbf{Q}_p)$ et $(\varphi, \Gamma)$-modules", pp. 281–509 in *Représentations p-adiques de groupes p-adiques*, vol. II: Représentations de $\mathrm{GL}_2(\mathbf{Q}_p)$ et $(\varphi, \Gamma)$-modules, edited by L. Berger et al., Astérisque **330**, Société Mathématique de France, Paris, 2010. MR Zbl

[Colmez et al. 2014] P. Colmez, G. Dospinescu, and V. Paškūnas, "The $p$-adic local Langlands correspondence for $\mathrm{GL}_2(\mathbb{Q}_p)$", *Camb. J. Math.* **2**:1 (2014), 1–47. MR Zbl

[EGA I 1971] A. Grothendieck and J. A. Dieudonné, *Eléments de géométrie algébrique, I*, Grundl. Math. Wissen. **166**, Springer, 1971. MR Zbl

[EGA $\mathrm{III}_2$ 1963] A. Grothendieck, "Eléments de géométrie algébrique, III: Étude cohomologique des faisceaux cohérents, II", *Inst. Hautes Études Sci. Publ. Math.* **17** (1963), 5–91. MR Zbl

[EGA $\mathrm{IV}_2$ 1965] A. Grothendieck, "Éléments de géométrie algébrique. IV: Étude locale des schémas et des morphismes de schémas. II", *Inst. Hautes Études Sci. Publ. Math.* **24** (1965), 5–231. MR

[EGA $\mathrm{IV}_3$ 1966] A. Grothendieck, "Eléments de géométrie algébrique, IV: Étude locale des schémas et des morphismes de schémas, III", *Inst. Hautes Études Sci. Publ. Math.* **28** (1966), 5–255. MR Zbl

[EGA $\mathrm{IV}_4$ 1967] A. Grothendieck, "Eléments de géométrie algébrique, IV: Étude locale des schémas et des morphismes de schémas, IV", *Inst. Hautes Études Sci. Publ. Math.* **32** (1967), 5–361. MR Zbl

[Eisenbud 1995] D. Eisenbud, *Commutative algebra: with a view toward algebraic geometry*, Graduate Texts in Mathematics **150**, Springer, 1995. MR Zbl

[Fargues and Fontaine 2018] L. Fargues and J.-M. Fontaine, *Courbes et fibrés vectoriels en théorie de Hodge p-adique*, Astérisque **406**, Société Mathématique de France, Paris, 2018. MR Zbl

[Fontaine and Rapoport 2005] J.-M. Fontaine and M. Rapoport, "Existence de filtrations admissibles sur des isocristaux", *Bull. Soc. Math. France* **133**:1 (2005), 73–86. MR Zbl

[Genestier and Lafforgue 2011] A. Genestier and V. Lafforgue, "Théorie de Fontaine en égales caractéristiques", *Ann. Sci. Éc. Norm. Supér.* (4) **44**:2 (2011), 263–360. MR Zbl

[Genestier and Lafforgue 2012] A. Genestier and V. Lafforgue, "Structures de Hodge–Pink pour les $\varphi/\mathfrak{S}$-modules de Breuil et Kisin", *Compos. Math.* **148**:3 (2012), 751–789. MR Zbl

[Grothendieck 1962] A. Grothendieck, *Fondements de la géométrie algébrique*, Secrétariat mathématique, Paris, 1962. MR Zbl

[Hartl 2011] U. Hartl, "Period spaces for Hodge structures in equal characteristic", *Ann. of Math.* (2) **173**:3 (2011), 1241–1358. MR Zbl

[Hartl and Viehmann 2011] U. Hartl and E. Viehmann, "The Newton stratification on deformations of local $G$-shtukas", *J. Reine Angew. Math.* **656** (2011), 87–129. MR Zbl

[Hellmann 2011] E. Hellmann, "On families of weakly admissible filtered $\varphi$-modules and the adjoint quotient of $\mathrm{GL}_d$", *Doc. Math.* **16** (2011), 969–991. MR Zbl

[Hellmann 2013] E. Hellmann, "On arithmetic families of filtered $\varphi$-modules and crystalline representations", *J. Inst. Math. Jussieu* **12**:4 (2013), 677–726. MR Zbl

[Huber 1993] R. Huber, "Continuous valuations", *Math. Z.* **212**:3 (1993), 455–477. MR Zbl

[Huber 1994] R. Huber, "A generalization of formal schemes and rigid analytic varieties", *Math. Z.* **217**:4 (1994), 513–551. MR Zbl

[Huber 1996] R. Huber, *Étale cohomology of rigid analytic varieties and adic spaces*, Aspects of Mathematics **E30**, Vieweg, Braunschweig, 1996. MR Zbl

[Kedlaya 2008] K. S. Kedlaya, "Slope filtrations for relative Frobenius", pp. 259–301 in *Représentations $p$-adiques de groupes $p$-adiques*, vol. I: Représentations galoisiennes et $(\varphi, \Gamma)$-modules, Astérisque **319**, Société Mathématique de France, Paris, 2008. MR Zbl

[Kedlaya et al. 2014] K. S. Kedlaya, J. Pottharst, and L. Xiao, "Cohomology of arithmetic families of $(\varphi, \Gamma)$-modules", *J. Amer. Math. Soc.* **27**:4 (2014), 1043–1115. MR Zbl

[Kisin 2006] M. Kisin, "Crystalline representations and $F$-crystals", pp. 459–496 in *Algebraic geometry and number theory*, edited by V. Ginzburg, Progr. Math. **253**, Birkhäuser, Boston, 2006. MR Zbl

[Kisin 2008] M. Kisin, "Potentially semi-stable deformation rings", *J. Amer. Math. Soc.* **21**:2 (2008), 513–546. MR Zbl

[Kisin 2009] M. Kisin, "The Fontaine–Mazur conjecture for $\mathrm{GL}_2$", *J. Amer. Math. Soc.* **22**:3 (2009), 641–690. MR Zbl

[Kottwitz 2006] R. E. Kottwitz, "Dimensions of Newton strata in the adjoint quotient of reductive groups", *Pure Appl. Math. Q.* **2**:3 (2006), 817–836. MR Zbl

[Laszlo and Sorger 1997] Y. Laszlo and C. Sorger, "The line bundles on the moduli of parabolic $G$-bundles over curves and their sections", *Ann. Sci. École Norm. Sup.* (4) **30**:4 (1997), 499–525. MR Zbl

[Laumon and Moret-Bailly 2000] G. Laumon and L. Moret-Bailly, *Champs algébriques*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) **39**, Springer, 2000. MR Zbl

[Lazard 1962] M. Lazard, "Les zéros des fonctions analytiques d'une variable sur un corps valué complet", *Inst. Hautes Études Sci. Publ. Math.* **14** (1962), 47–75. MR Zbl

[Lütkebohmert 1977] W. Lütkebohmert, "Vektorraumbündel über nichtarchimedischen holomorphen Räumen", *Math. Z.* **152**:2 (1977), 127–143. MR Zbl

[Mumford 1999] D. Mumford, *The red book of varieties and schemes*, expanded ed., Lecture Notes in Mathematics **1358**, Springer, 1999. MR Zbl

[Pappas and Rapoport 2009] G. Pappas and M. Rapoport, "Φ-modules and coefficient spaces", *Mosc. Math. J.* **9**:3 (2009), 625–663. MR Zbl

[Paškūnas 2013] V. Paškūnas, "The image of Colmez's Montreal functor", *Publ. Math. Inst. Hautes Études Sci.* **118** (2013), 1–191. MR Zbl

[Pink 1997] R. Pink, "Hodge structures over function fields", preprint, 1997, available at http://www.math.ethz.ch/~pinkri/ftp/HS.pdf.

[Rotman 2009] J. J. Rotman, *An introduction to homological algebra*, 2nd ed., Springer, 2009. MR Zbl

[Schneider and Teitelbaum 2006] P. Schneider and J. Teitelbaum, "Banach–Hecke algebras and $p$-adic Galois representations", *Doc. Math.* Extra Vol. (2006), 631–684. MR Zbl

[Serre 1979] J.-P. Serre, *Local fields*, Graduate Texts in Mathematics **67**, Springer, 1979. MR Zbl

[Sorensen 2013] C. M. Sorensen, "A proof of the Breuil–Schneider conjecture in the indecomposable case", *Ann. of Math.* (2) **177**:1 (2013), 367–382. MR Zbl

[Varshavsky 2004] Y. Varshavsky, "Moduli spaces of principal $F$-bundles", *Selecta Math.* (*N.S.*) **10**:1 (2004), 131–166. MR Zbl

[Wang-Erickson 2018] C. Wang-Erickson, "Algebraic families of Galois representations and potentially semi-stable pseudodeformation rings", *Math. Ann.* **371**:3-4 (2018), 1615–1681. MR Zbl

Communicated by Brian Conrad
Received 2015-10-08 Revised 2019-05-29 Accepted 2019-11-24

[Urs Hartl] *Mathematisches Institut, Universität Münster, Münster, Germany*

e.hellmann@uni-muenster.de *Mathematisches Institut, Universität Münster, Münster, Germany*

msp

ALGEBRA AND NUMBER THEORY 14:5 (2020)
https://doi.org/10.2140/ant.2020.14.1123

# On the group of purely inseparable points of an abelian variety defined over a function field of positive characteristic, II

Damian Rössler

Let $A$ be an abelian variety over the function field $K$ of a curve over a finite field. We describe several mild geometric conditions ensuring that the group $A(K^{\mathrm{perf}})$ is finitely generated and that the $p$-primary torsion subgroup of $A(K^{\mathrm{sep}})$ is finite. This gives partial answers to questions of Scanlon, Ghioca and Moosa, and Poonen and Voloch. We also describe a simple theory (used to prove our results) relating the Harder–Narasimhan filtration of vector bundles to the structure of finite flat group schemes of height one over projective curves over perfect fields. Finally, we use our results to give a complete proof of a conjecture of Esnault and Langer on Verschiebung divisibility of points in abelian varieties over function fields.



## 1. Introduction

Let $k$ be a finite field characteristic $p > 0$ and let $S$ be a smooth, projective and geometrically connected curve over $k$. Let $K := \kappa(S)$ be its function field. Let $A$ be an abelian variety of dimension $g$ over $K$. Choose an algebraic closure $\overline{K}$ of $K$. Let $K^{\mathrm{perf}} \subseteq \overline{K}$ be the maximal purely inseparable extension of $K$, let $K^{\mathrm{sep}} \subseteq \overline{K}$ be the maximal separable extension of $K$ and let $K^{\mathrm{unr}} \subseteq K^{\mathrm{sep}}$ be the maximal separable

*MSC2010:* primary 11J95; secondary 11G10, 14G25.
*Keywords:* abelian varieties, rational points, purely inseparable extensions, Frobenius, Verschiebung.

extension of $K$, which is unramified above every place of $K$. Finally, we let $\mathcal{A}$ be a smooth commutative group scheme over $S$ such that $\mathcal{A}_K = A$. We shall write $\omega_{\mathcal{A}} := \epsilon^*_{\mathcal{A}/S}(\Omega_{\mathcal{A}/S})$ for the restriction of the cotangent sheaf of $\mathcal{A}$ over $S$ via the zero section $\epsilon_{\mathcal{A}/S} : S \to \mathcal{A}$ of $\mathcal{A}$. We shall say that $\omega_{\mathcal{A}}$ is the Hodge bundle of $\mathcal{A}$.

If $G$ is an abelian group, we shall write

$$\mathrm{Tor}_p(G) := \{x \in G \mid \exists n \geq 0 : p^n \cdot x = 0\} \quad \text{and} \quad \mathrm{Tor}^p(G) := \{x \in G \mid \exists n \geq 0 : n \cdot x = 0 \wedge (n, p) = 1\}.$$

The aim of this text is to prove the following two theorems and to give a proof of a conjecture of Esnault and Langer (see further below).

**Theorem 1.1.** (a) *Suppose that $A$ is geometrically simple. If $A(K^{\mathrm{perf}})$ is finitely generated and of rank $> 0$, then $\mathrm{Tor}_p(A(K^{\mathrm{sep}}))$ is a finite group.*

(b) *Suppose that $A$ is an ordinary (not necessarily simple) abelian variety. If $\mathrm{Tor}_p(A(K^{\mathrm{sep}}))$ is a finite group, then $A(K^{\mathrm{perf}})$ is finitely generated.*

**Theorem 1.2.** *Suppose that $\mathcal{A}$ is a semiabelian scheme and that $A$ is a geometrically simple abelian variety over $K$. If $\mathrm{Tor}_p(A(K^{\mathrm{sep}}))$ is infinite, then*

(a) *$\mathcal{A}$ is an abelian scheme;*

(b) *there is $r_A \geq 0$ such that $p^{r_A} \cdot \mathrm{Tor}_p(A(K^{\mathrm{sep}})) \subseteq \mathrm{Tor}_p(A(K^{\mathrm{unr}}))$.*

*Furthermore, there is*

(c) *an abelian scheme $\mathcal{B}$ over $S$;*

(d) *an $S$-isogeny $\mathcal{A} \to \mathcal{B}$, whose degree is a power of $p$ and such that the corresponding isogeny $\mathcal{A}_K \to \mathcal{B}_K$ is étale;*

(e) *an étale $S$-isogeny $\mathcal{B} \to \mathcal{B}$ whose degree is $> 1$ and is a power of $p$,*

*and*

(f) (Voloch) *if $A$ is ordinary then the Kodaira–Spencer rank of $A$ is not maximal;*

(g) *if $\dim(A) \leqslant 2$ then $\mathrm{Tr}_{\overline{K}|\bar{k}}(A_{\overline{K}}) \neq 0$;*

(h) *for all closed points $s \in S$, the $p$-rank of $\mathcal{A}_s$ is $> 0$.*

Here $\mathrm{Tr}_{\overline{K}|\bar{k}}(A_{\overline{K}})$ is the $\overline{K}|\bar{k}$-trace of $A_{\overline{K}}$. This is an abelian variety over $\bar{k}$. See Section 9A.

Theorems 1.1 and 1.2 (b) have applications in the context of the work of Poonen and Voloch on the Brauer–Manin obstruction over function fields. In particular Theorems 1.1 and 1.2 (b) show that the conclusion of [Poonen and Voloch 2010, Theorem B] holds whenever the underlying abelian variety is geometrically simple, has semistable reduction and violates any of the conditions in Theorem 1.2, in particular if it has a point of bad reduction. Theorems 1.1 and 1.2 (b) also feed into the "full" Mordell–Lang conjecture. See [Scanlon 2005, after Claim 4.4; Abramovich and Voloch 1992, Introduction] for this conjecture. In particular, in conjunction with the main result of [Ghioca and Moosa 2006], Theorems 1.1

and 1.2 (b) show that the "full" Mordell–Lang conjecture holds if the underlying abelian variety is ordinary, geometrically simple, has semistable reduction and violates any of the conditions in Theorem 1.2, in particular if it has a point of bad reduction.

Let now $L$ be a field, which is finitely generated as a field over an algebraically closed field $l_0$ of characteristic $p$. Let $C$ be an abelian variety over $L$.

**Conjecture 1.3** (Esnault-Langer). *Suppose that for all $\ell \geqslant 0$ we are given a point $x_\ell \in C^{(p^\ell)}(L)$ and suppose that for all $\ell \geqslant 1$, we have $V_{C^{(p^\ell)}/L}(x_\ell) = x_{\ell-1}$. Then the image of $x_0$ in $C(L)/\mathrm{Tr}_{L|l_0}(C)(l_0)$ is a torsion point, which is of order prime to $p$.*

See [Esnault and Langer 2013, Remark 6.3 and after Lemma 6.5]. This conjecture is important in the theory of stratified bundles in positive characteristic; see [Esnault and Langer 2013, Question 3 in the introduction] for details.

Here $C^{(p^\ell)}$ is the base change of $C$ by the $\ell$-th power of the absolute Frobenius morphism on Spec $L$ and $V_{C^{(p^\ell)}/L} : C^{(p^\ell)} \to C^{(p^{\ell-1})}$ is the Verschiebung morphism. The abelian variety $\mathrm{Tr}_{L|l_0}(C)$ is the $L|l_0$-trace of $C$ (see Section 9A). It is an abelian variety over $l_0$ and the variety $\mathrm{Tr}_{L|l_0}(C)_L$ comes with an injective morphism to $C$. This gives in particular an injective map $\mathrm{Tr}_{L|l_0}(C)(l_0) \to C(L)$. The Lang–Néron theorem (see [Lang 1983, Chapter 6, Theorem 2]) asserts that $C(L)/\mathrm{Tr}_{L|l_0}(C)(l_0)$ is a finitely generated group. Thus $\mathrm{Tr}_{L|l_0}(C)(l_0) \subseteq C(L)$ is precisely the subgroup of $C(L)$ consisting of divisible elements (i.e., elements divisible by any integer).

In the present text, we shall call a point $x_0 \in C(L)$ with the property described in Conjecture 1.3 an *indefinitely Verschiebung divisible point.* We shall write $\mathrm{IVD}(C) = \mathrm{IVD}(C, L) \subseteq C(L)$ for the subgroup of indefinitely Verschiebung divisible points.

We prove:

**Theorem 1.4.** *Conjecture 1.3 holds.*

Note that Theorem 1.4 has the following consequence, which is of independent interest: if $C$ is as in Conjecture 1.3, $C$ is ordinary and $\mathrm{Tr}_{L^{\mathrm{perf}}|l_0}(C_{L^{\mathrm{perf}}}) = 0$ then

$$\bigcap_{j\geq 0} p^j \cdot C(L^{\mathrm{perf}}) = \mathrm{Tor}^p(C(L^{\mathrm{perf}})).$$

To see this, let $x \in C(L^{\mathrm{perf}})$. Let $L_1|L$ be a finite purely inseparable extension, which is a field of definition for $x$. Remember that the multiplication by $p$ endomorphism of $C$ is the composition of the Verschiebung morphism with the relative Frobenius morphism, which is purely inseparable. Also, recall that since $C$ is ordinary, the Verschiebung morphism is (by definition) separable. Note finally that since $\mathrm{Tr}_{L^{\mathrm{perf}}|l_0}(C_{L^{\mathrm{perf}}}) = 0$ we also have $\mathrm{Tr}_{L_1|l_0}(C_{L_1}) = 0$. In particular, if $x \in \bigcap_{j\geq 0} p^j \cdot C(L^{\mathrm{perf}})$ then $x$ is an indefinitely Verschiebung divisible element of $C(L_1)$ and thus must lie in $\mathrm{Tor}^p(C(L_1)) \subseteq \mathrm{Tor}^p(C(L^{\mathrm{perf}}))$ according to Theorem 1.4. The inclusion $\mathrm{Tor}^p(C(L^{\mathrm{perf}})) \subseteq \bigcap_{j\geq 0} p^j \cdot C(L^{\mathrm{perf}})$ is straightforward.

***Outline of the paper.*** The basic strategy of the paper hinges on Lemma 4.8 below. This Lemma associates a maximal multiplicative subgroup scheme with any finite flat group scheme of height one over $S$. The existence of this subgroup scheme is not straightforward and follows from an analysis of the Harder–Narasimhan filtration of (a Frobenius twist of) the coLie algebra of the group scheme. This analysis is carried out in Section 4B.

One can apply Lemma 4.8 to the kernel of the relative Frobenius morphism $F_{\mathcal{A}/S}: \mathcal{A} \to \mathcal{A}^{(p)}$, replace $\mathcal{A}$ by the resulting quotient and repeat this construction ad infinitum, stopping only when the maximal multiplicative subgroup scheme is trivial.

It is then a basic (unresolved) question to determine minimal geometric conditions on $\mathcal{A}$ ensuring that the resulting sequence of semiabelian schemes stops. This also makes sense (and seems important to us) if $k$ is replaced by any perfect field of characteristic $p > 0$ (not only when $k$ is finite).

This question turns out to be intimately related to Theorems 1.1, 1.2 and 1.4. To explain why, we shall first quote a result, which improves on (and elucidates) Lemma B.2 in the Appendix. This result is proven in [Rössler 2019a], which builds on the present article. We shall only need Lemma B.2 in the present text but for conceptual clarity, we shall present the improved result in this outline. Let $E \subseteq S$ be the finite set of points $s \in S$ where $\mathcal{A}_s$ is not an abelian variety. Let $U := S \backslash E$. We first recall a classical result:

**Theorem 1.5** (Artin–Milne). *There is a canonical injective group homomorphism*

$$A^{(p)}(K)/F_{A/K}(A(K)) \hookrightarrow \mathrm{Hom}_K(F_K^*(\omega_K), \Omega_{K/k}).$$

Here $F_K$ is the absolute Frobenius endomorphism of $K$ (the $p$-th power map). See [Artin and Milne 1976, III.3.5.6] for the proof, which works in a more general setting. In [Rössler 2019a] this is refined as follows:

**Theorem 1.6** (R.). *The image of the Artin–Milne map lies inside the subgroup* $\mathrm{Hom}_C(F_S^*(\omega), \Omega_{S/k}(E))$ *of* $\mathrm{Hom}_K(F_K^*(\omega_K), \Omega_{K/k})$.

Here $F_S$ is the absolute Frobenius endomorphism of $S$. Here we write $\Omega_{S/k}(E) := \Omega_{S/k}(E) \otimes \mathcal{O}_S(E)$ and $E$ is understood as a divisor with no multiplicities. Theorem 1.6 refines Lemma B.2 below (for the knowledgeable reader, in [Rössler 2019a] it is even proven that the image of the Selmer group of the relative Frobenius morphism lies in $\mathrm{Hom}_C(F_S^*(\omega), \Omega_{S/k}(E))$). The group $\mathrm{Hom}_C(F_S^*(\omega), \Omega_{S/k}(E))$ can be understood as the target of an Abel–Jacobi map in logarithmic Higgs cohomology, although to give a precise meaning to this interpretation would require the development of a good theory of Higgs bundles in positive characteristic (which does not exist at the moment, to the author's knowledge). This theorem is proven by providing a geometric interpretation for the Artin–Milne map and analysing its poles, making essential use of Faltings and Chai's semistable compactification of the universal abelian scheme. The existence of this compactification allows us to show that the poles are at most logarithmic, which is in essence the content of Theorem 1.6. Let us now explain why Theorem 1.6 is relevant for Theorem 1.1.

Consider, e.g., (b) in Theorem 1.1. Suppose that $A(K^{\mathrm{perf}})$ is not finitely generated. We have

$$A(K^{\mathrm{perf}}) = \bigcup_{i \geq 0} A(K^{p^{-i}})$$

and by the Lang–Néron theorem (see also Section 9A) $A(K^{p^{-i}})$ is finitely generated. Hence for infinitely many $i \geq 0$, we must have

$$A^{(p^{i+1})}(K)/F_{A^{(p^i)}/K}(A(K)) \simeq A(K^{p^{-i-1}})/A(K^{p^{-i}}) \neq 0.$$

In particular, for infinitely many $i \geq 0$, we must have

$$\mathrm{Hom}_C(F_S^{\circ(i+1),*}(\omega), \Omega_{S/k}(E)) \neq 0$$

according to Theorem 1.6. If now the vector bundle $\omega$ were ample, this would lead to a contradiction, because if $i$ is large enough and $\omega$ is ample then there cannot be any morphism from $F_S^{\circ(i+1),*}(\omega)$ to $\Omega_{S/k}(E)$. This was already noticed in the earlier article [Rössler 2015], where details are given. One can refine this line of reasoning as follows. If $\omega$ is not ample and $A$ is ordinary then one can show that $\omega$ must have a certain nontrivial quotient, which is semistable of degree 0. This nontrivial quotient turns out to be induced by the maximal multiplicative subgroup scheme mentioned above. Calling it $G_{\mathcal{A}}$, we may then replace $\mathcal{A}$ by $\mathcal{A}/G_{\mathcal{A}}$. The group $(\mathcal{A}/G_{\mathcal{A}})_K(K^{\mathrm{perf}})$ will again be infinitely generated, since the morphism $A \to (\mathcal{A}/G_{\mathcal{A}})_K$ has finite kernel. Hence we can repeat the above reasoning for $\mathcal{A}/G_{\mathcal{A}}$ and we obtain an infinite sequence of isogenous abelian varieties. The next step in the proof of Theorem 1.1 (b) is to show that in this sequence, there are finitely many isomorphism classes. This follows from the fact that the degrees of $\omega_{\mathcal{A}}$ and $\mathcal{A}/G_{\mathcal{A}}$ are the same and more generally the degrees of the Hodge bundles of all the semiabelian schemes in the sequence are the same. This is a consequence of a computation involving the cotangent complex of the quotient morphism (see Lemma 4.12). It then follows from a classical reasoning involving moduli spaces of abelian varieties, familiar from Zarhin's proof of the Tate conjecture over function fields, that the sequence contains only finitely many isomorphism classes. We can thus conclude that, up to isogeny, $A$ contains a nontrivial finite endomorphism, whose kernel is multiplicative. The dual of this endomorphism is then separable and this shows that $\mathrm{Tor}_p(A^\vee)(K^{\mathrm{sep}})$ is infinite (consider the kernels of its powers). Since $A^\vee$ is isogenous to $A$, we see that $\mathrm{Tor}_p(A)(K^{\mathrm{sep}})$ is also infinite. This concludes our outline of the proof of Theorem 1.1 (b).

For Theorem 1.1 (a), we consider the quotients of $A$ by finite subgroups of $\mathrm{Tor}_p(A)(K^{\mathrm{sep}})$ of increasing size. These quotients also run through finitely many isomorphism classes by a similar reasoning and we thus see that if $\mathrm{Tor}_p(A)(K^{\mathrm{sep}})$ is infinite then, up to isogeny, $A$ is endowed with a separable finite endomorphism. The dual of this endomorphism is then purely inseparable and of degree a positive power of $p$, and if $A^\vee(K)$ is not finite, we may show that $A^\vee(K^{\mathrm{perf}})$ is infinitely generated by considering the inverse images of $A(K)$ under the powers of this endomorphism. If now $A^\vee(K^{\mathrm{perf}})$ is not finitely generated, neither is $A(K^{\mathrm{perf}})$, since $A$ and $A^\vee$ are isogenous. This concludes our outline of the proof of Theorem 1.1 (a).

In Theorem 1.2, we start out as in Theorem 1.1 (a) and we again obtain, up to isogeny, a separable finite endomorphism of degree a positive power of $p$. The rest of the theorem investigates the geometric

consequences of the existence of this endomorphism. The most interesting consequence is the fact that it implies that $\mathcal{A}$ must be an abelian scheme (if $A$ is geometrically simple). This is (a) in Theorem 1.2. The main point here is that the endomorphism extends to an étale endomorphism of $\mathcal{A}$. If $\mathcal{A}$ had a fibre with a toric part then the endomorphism would induce an automorphism of the toric part, because tori only have infinitesimal $p$-primary subgroups in characteristic $p$ and these are only étale if they are trivial. This fact forces the whole endomorphism to be an automorphism, which is impossible. The proof of (c), (d) and (e) are straightforward and not much more than a rewording of the fact that there are only finitely many isomorphism classes in the set of quotients described above. The proof of (b) follows essentially from a variant of the fact that, under (a), the above endomorphism extends to an everywhere étale and finite endomorphism of $\mathcal{A}$. This also easily gives a proof of (h). The proof of (g) is based on class field theory and the Serre–Tate theory of canonical liftings. First, up to a finite extension, the field extension generated by the points of $\mathrm{Tor}_p(A)(K^{\mathrm{sep}})$ is everywhere unramified by (a) and (b). If $\mathrm{Tor}_p(A)(K^{\mathrm{sep}}) = \mathrm{Tor}_p(A)(\overline{K})$ then a simple application of the Serre–Tate theory of canonical liftings shows that $A_{\overline{K}}$ is the base change of an abelian variety defined over $\bar{k}$. Hence it must be contained in the Hilbert class field of $K$, which is but a constant field extension (i.e., comes from an extension of $k$), up to a finite extension. So if $\mathrm{Tor}_p(A)(K^{\mathrm{sep}})$ is infinite then it is an infinite torsion subset of $A(K\bar{k})$, which is finitely generated by the Lang–Néron theorem if the trace of $A$ vanishes. This is a contradiction.

We now turn to Theorem 1.4. Esnault and Langer [2013, Theorem 6.2], using a height argument due to Raynaud, proved that the image of $x_0$ in $C(L)/\mathrm{Tr}_{L|l_0}(C)(l_0)$ is a torsion point under the assumption that $C$ has everywhere potential good reduction in codimension one. Their argument works as follows. Choose a polarisation on $C$. This induces polarisations on all the $C^{(p^\ell)}$ by base change. A simple computation shows that if a point $x \in C(L)$ has a preimage $y$ in $C^{(p)}(K)$ under the Verschiebung map then the height of $x$ with respect to the polarisation is $p$ times the height of $y$ with respect to the base changed polarisation. Now if $C$ has everywhere good reduction in codimension one, there is an abelian scheme $\mathcal{C}$ extending $C$ on an open subset with complement of codimension $\geq 2$ of a normal complete model $V$ of $L$ and the polarisations on $C$ and $C^{(p)}$ naturally extend to this open subset. This implies that the heights of $x$ are $y$ (with respect to the polarisations and a choice of ample line bundle on $V$) are integers, because they can then computed in a completely geometric fashion. In particular, the height of $x$ is an integer divisible by $p$. Repeating this argument with $y$, one sees that the height of $x$ is divisible by arbitrarily high powers of $p$ and one concludes that it must vanish. Then the conclusion follows from a theorem of Lang (see [Conrad 2006, Theorem 9.15]). The argument described above breaks down in the presence of bad reduction in codimension one because the orders of the component groups of the special fibres of the local Néron models of the varieties $C^{(p^\ell)}$ increase with $\ell$ if they are not trivial and this introduces denominators in the heights.

Our approach to Theorem 1.4 is again via the infinite sequence of quotients described at the beginning of the outline. This sequence will effectively replace the sequence of the $C^{(p^\ell)}$. It has the advantage over the sequence of the $C^{(p^\ell)}$ that it falls inside a bounded family of abelian varieties (see below), making it possible to control the order of the (analogues of the) images of the $x_\ell$ in the component groups of the Néron models. This makes a similar height computation possible. The proof is in several steps.

Step (0). Reduction to the case where $L$ is the function field of a smooth and projective curve $B$ over $l_0$. This follows from a Bertini type argument — see Appendix C.

Step (1). We consider the images of the $x_\ell$ under the Artin–Milne map. A crucial point is that these images must be compatible under the Verschiebung morphisms (see diagram (8) below) and this constrains the image of $x_1$ under the Artin–Milne map. Using Lemma B.2 (or Theorem 1.6), the theory of semistable sheaves in positive characteristic and various global results on finite flat group schemes of height one in a global situation proven in Section 4, we show that the image of $x_1$ under the Artin–Milne map must factor through the coLie algebra of the maximal multiplicative subgroup $(\ker F_{\mathcal{C}/B})_\mu$ of $\ker F_{\mathcal{C}/B}$. This implies that the image of $x_1$ in $(C^{(p)}/(\ker F_{\mathcal{C}/B})^{(p)}_{\mu,L})(L) = (C^{(p)}/G_{\mathcal{C}^{(p)},L})(L)$ maps to 0 under the Artin–Milne map. From the definitions, this means that the image of $x_0$ in $(C/G_{\mathcal{C},L})(L)$ is divisible by $p$ in $(C/G_{\mathcal{C},L})(L)$. Suppose for simplicity that $C$ has a semiabelian model $\mathcal{C}$ over $B$. We can now repeat this process and we obtain a sequence of purely inseparable morphisms $\psi_i : \mathcal{C} \to \mathcal{C}_i$ of increasing degree, such that $\psi_{i,L}(x_0)$ in $C_i$ is divisible by $p^i$ in $C_i(L)$.

Step (2). We choose a polarisation $\phi_{D_0} : C \to C^\vee$. The image of $x_0$ under $\phi_{D_0}$ is of course also indefinitely Verschiebung divisible. We identify $\phi_{D_0}(x_0)$ with a line bundle $M$ on $C$. Since $\phi_{D_0}(x_0)$ is indefinitely Verschiebung divisible, there are line bundles $M_i$ on $C^{(p^i)}$ such that $M$ is the pull-back of $M_i$ by the morphism $C \to C^{(p^i)}$ arising by composing relative Frobenii. The morphism $C \to C^{(p^i)}$ factors through $\psi_{i,L}$ by construction. Hence there are line bundles $J_i$ on the $C_i$ such that $\psi^*_{i,L}(J_i) = M$.

Step (3). We now compute the height pairing between $x_0$ and $M$. This can easily be seen to equal the height pairing between $\psi_{i,L}(x_0)$ and $J_i$. Since $\psi_{i,L}(x_0)$ is divisible by $p^i$, we see that the height pairing between $x_0$ and $M$ is divisible by $p^i$. If the $\mathcal{C}_i$ were all abelian schemes we could deduce (like Raynaud–Esnault–Langer above) that the height pairing between $x_0$ and $M$ must vanish, because then all the values of the various height pairing would be integral. However, we cannot assume this.

Step (4). All the $C_i$ are essentially part of a bounded family of abelian varieties over $L$ because the degrees of the Hodge bundles of the $\mathcal{C}_i$ are all equal (see above in the outline). Using this, one can prove that there is an infinite set $I_0 \subseteq \mathbb{N}$ such that if $i \in I_0$ the image of any element of $C_i(L)$ in the component groups of the Néron model of $C_i$ has an order, which is bounded independently of $i$. This follows from Proposition A.2 (a) in the Appendix. The gist of the argument is that in a bounded family of semiabelian varieties over $B$, it is possible to smoothly compactify the generic fibre, up to normalisation in a finite extension of the function field of the parameter space. This would follow from resolution of singularities but in the present situation is a consequence of the work of Mumford, Chai–Faltings and Künnemann (see [Künnemann 1998, Theorem 4.2]). This means that the abelian varieties in the family almost all have regular compactifications with a bounded number of geometric fibres over $B$. This bound is also a bound for the order of the image of a rational point in the component groups of the Néron model.

Step (5). In view of Step (4), if we replace $x_0$ by a certain multiple of $x_0$, all the height pairing in sight are integers. Hence the divisibility argument envisaged in Step (3) can be carried out and yields that the

height pairing of $x_0$ and $M$ vanishes. This pairing is by construction twice the Néron–Tate height of $x_0$ with respect to the polarisation $\phi_{D_0}$ and we conclude from a theorem of Lang [1983] that the image of $x_0$ in $C(L)/\mathrm{Tr}_{L|l_0}(C)(l_0)$ is a torsion point. It remains to show that its order is prime to $p$.

Step (6). We first show that we may suppose that $\mathrm{Tr}_{\bar{L}|l_0}(C_{\bar{L}}) = 0$. This is not completely straightforward, because when one passes to a finite extension in Conjecture 1.3, one loses control of part of the torsion of $C(L)/\mathrm{Tr}_{L|l_0}(C)(l_0)$. However, although the parasitical torsion subgroup that might appear is not known, its exponent only depends on the degree of the extension. This degree can be taken to be the same for all the Frobenius twists of $C$ and the information one gathers from this suffices to prove the conjecture, provided one can prove it for a finite extension. Thus we may suppose that $\dim(\mathrm{Tr}_{\bar{L}|l_0}(C_{\bar{L}})) = \dim(\mathrm{Tr}_{L|l_0}(C))$ and then, after quotienting by $\mathrm{Tr}_{L|l_0}(C)$, that $\mathrm{Tr}_{\bar{L}|l_0}(C_{\bar{L}}) = 0$. Now recall that the $C_i$ are essentially part of a bounded family of abelian varieties over $L$ (see step (4)). Using this, and the fact that now $\mathrm{Tr}_{\bar{L}|l_0}(C_{i,\bar{L}}) = 0$ for all $i \geq 1$, one can prove that there is an infinite set $I_0 \subseteq \mathbb{N}$ such that if $i \in I_0$, the cardinality of the torsion subgroup of $C_i(L)$ is uniformly bounded. This follows from Proposition A.2 (b) in the Appendix. To finish the proof of Theorem 1.4, suppose that $x_0$ is a nonzero torsion point, which is indefinitely Verschiebung divisible. Since the image of $x_0$ in $C_i(L)$ is divisible by $p^i$, we see that the torsion group of $C_i(L)$ has an element of order $p^{i+1}$. This contradicts the above uniformity statement and shows that the order of $x_0$ must be prime to $p$.

The argument to prove the uniformity statement alluded to in Step (6) goes roughly as follows. One first notices that the torsion subgroup of a trace free abelian variety coincides with the set of elements of vanishing Néron–Tate height by the already quoted theorem of Lang. Thus they can be described as the points of a moduli space of sections, which is of finite type over $l_0$, at least for those torsion points, whose image in the component groups of the Néron model of the abelian variety is trivial. Since the abelian variety is trace free, the torsion subgroup is finite and thus this moduli space is finite. Using the uniformity statement in Step (4), we may assume that the torsion points of the $C_i(L)$ have trivial images in the components of the corresponding Néron models, up to multiplication by a fixed integer (independent of $i$ running through an infinite set). The number of irreducible components of the moduli space of each $C_i$ is now uniformly bounded, since the $C_i$ are part of a bounded family. This gives a uniform bound for the torsion subgroups of the $C_i(L)$.

The reader may enjoy the talk [Rössler 2019b] as an introduction to parts of the present article.

The structure of the article is as follows. In Section 2, we state various intermediate results, from which we shall deduce Theorems 1.1 and 1.2. Theorem 2.1 in Section 2A is of independent interest and is (we feel) likely to be useful for the study of the geometry of (especially ordinary) abelian varieties in general. The results in Section 2A are deduced from some results in the theory of finite flat groups schemes of height one over $S$, most of which follow from the existence of a Harder–Narasimhan filtration on their Lie algebras. These results on finite flat group schemes are proven in Section 4 and for the convenience of the reader, we included a section (Section 3) listing the results on semistable sheaves over curves in positive characteristic that we need. To the knowledge of the author, there are very few general results

on the structure of finite flat group schemes in a global situation (e.g., when the base is not affine) and it seems that it is the first time that the theory of semistability of vector bundles is being used in this context. In [Bost 2004] a similar idea is used in characteristic 0, where it is applied to the study of formal groups over curves (recall that all groups schemes are smooth in characteristic 0, so the Lie algebras of finite flat group schemes vanish in characteristic 0). Lemma 4.4 below (which concerns finite flat group schemes of height one) is inspired by [Bost 2004, Lemma 2.9]. A prototype of Lemma 4.4 can be found in [Shepherd-Barron 1992, Lemma 9.1.3.1] but it is not applied to the study of group schemes there. The key results here are the Lemmata 4.4 and 4.8, which will hopefully lead to further generalisations (e.g., in the situation when the base scheme is of dimension higher than one - in this direction, see [Langer 2015, Theorem 7.3]). The results in Section 2B do not require the theory of semistable sheaves and are based on geometric class field theory, the theory of Serre–Tate canonical liftings and on the existence of moduli schemes for abelian varieties. In Section 5, we prove the various claims made in Section 2A and in Section 6 we prove the claims made in Section 2B. In Section 7, we prove Theorem 1.1 and in Section 8 we prove Theorem 1.2. In Section 9B, we give a proof of Theorem 1.4. The proof of Theorem 1.4 is quite long and uses virtually all the other results proven in this text.

In his very interesting recent preprint, Xinyi Yuan [2018] uses some techniques which are also used in the present paper. They were discovered independently. His text focusses on the case where the base curve is the projective line. In particular, the "quotient process" used in step (2) of the proof of Theorem 1.4 and also in the proof of Theorem 1.2 also appears (over the projective line) in section 2.2 of [Yuan 2018]. Theorem 2.9 of [Yuan 2018] overlaps with the proof of Lemma 4.11.

The prerequisites for this article are algebraic geometry at the level of the EGA, familiarity with the basic theory of finite flat group schemes, as expounded in [Tate 1997] and a good knowledge of the theory of abelian schemes and varieties, as presented in [Milne 1986; Mumford 1970; Moret-Bailly 1985]. We also expect the reader to be familiar with the basic properties of Néron models (as in the chapter on basics of [Bosch et al. 1990]) and to have a working knowledge of Grothendieck topologies.

***Notation.*** If $X$ is an integral scheme, we write $\kappa(X)$ for the local ring at the generic point of $X$ (which is a field). If $X$ is a scheme of characteristic $p$, we denote the absolute Frobenius endomorphism of $X$ by $F_X$. If $f : X \to Y$ is a morphism between two schemes of characteristic $p$ and $\ell > 0$, abusing language, we denote by $X^{(p^\ell)}$ the fibre product of $f$ and $F_Y^{\circ\ell}$, where $F_Y^{\circ\ell}$ is the $\ell$-th power of the Frobenius endomorphism $F_Y$ of $Y$. If $G \to X$ is a group scheme, we write $\epsilon_{G/X} : X \to G$ for the zero section of $G$ and

$$\omega_{G/X} = \omega_G := \epsilon_{G/X}^*(\Omega_{G/X}).$$

If $X$ is of characteristic $p$, we shall write $F_{G/X} : G \to G^{(p)}$ for the relative Frobenius morphism. If in addition $G$ is flat and commutative, we shall write $V_{G^{(p)}/X} : G^{(p)} \to G$ for the corresponding Verschiebung morphism; we shall write

$$F_{G/X}^{(n)} : G \to G^{(p^n)} \quad \text{and} \quad V_{G^{(p^n)}/X}^{(j)} : G^{(p^n)} \to G^{(p^{n-j})}$$

for the compositions of morphisms

$$F_{G^{(p^{n-1})}/X} \circ \cdots \circ F_{G/X} \qquad \text{and} \qquad V_{G^{(p^{n-j+1})}/X} \circ V_{G^{(p^{n-j+2})}/X} \circ \cdots \circ V_{G^{(p^n)}/X}),$$

respectively. See [SGA $3_{\text{I}}$ 2011, Exposé $\text{VII}_{\text{A}}$, §4, "Frobeniuseries"] for the definition of the relative Frobenius morphism and the Verschiebung. If $G$ is finite flat and commutative, we shall write $G^\vee$ for the Cartier dual of $G$.

## 2. Intermediate results

We keep the notations and terminology of the introduction.

**2A.** ***Consequences of infinite generation of $A(K^{\text{perf}})$.*** We shall write

$$\overline{\text{rk}}_{\min}(\omega_{\mathcal{A}}) := \lim_{\ell\to\infty} \text{rk}((F_S^{\circ\ell,*}(\omega_{\mathcal{A}}))_{\min}) \quad \text{and} \quad \bar{\mu}_{\min}(\omega_{\mathcal{A}}) := \lim_{\ell\to\infty} \frac{\deg((F_S^{\circ\ell,*}(\omega_{\mathcal{A}}))_{\min})}{p^\ell \cdot \text{rk}((F_S^{\circ\ell,*}(\omega_{\mathcal{A}}))_{\min})}.$$

Here $F_S^{\circ\ell}$ is the $\ell$-th power of the absolute Frobenius endomorphism of $S$ and $(F_S^{\circ\ell,*}(\omega_{\mathcal{A}}))_{\min}$ is the semistable quotient with minimal slope of the vector bundle $F_S^{\circ\ell,*}(\omega_{\mathcal{A}})$. See Section 3 for details. Our main tool will be the following theorem.

**Theorem 2.1.** *There exists a* (*necessarily unique*) *multiplicative subgroup scheme $G_{\mathcal{A}} \hookrightarrow \ker F_{\mathcal{A}/S}$, with the following property*: *if $H$ is a finite, flat, multiplicative group scheme of height one over $S$ and $f : H \to \ker F_{\mathcal{A}/S}$ is a morphism of group schemes, then $f$ factors through $G_{\mathcal{A}}$.*

*If $A$ is ordinary and $\omega_{\mathcal{A}}$ is not ample then the order of $G_{\mathcal{A}}$ is $p^{\overline{\text{rk}}_{\min}(\omega_{\mathcal{A}})}$.*

*If $\phi : \mathcal{A} \to \mathcal{B}$ is a morphism of smooth commutative group schemes over $S$, then the restriction of $\phi$ to $G_{\mathcal{A}}$ factors through $G_{\mathcal{B}}$. Furthermore, we have* $\deg(\omega_{\mathcal{A}}) = \deg(\omega_{\mathcal{A}/G_{\mathcal{A}}})$.

Here $\mathcal{A}/G_{\mathcal{A}}$ is the "fppf quotient" of $\mathcal{A}$ by $G$, which is also a smooth commutative group scheme over $S$. See Proposition 4.1 below for details.

**Remark 2.2.** Note that $\bar{\mu}_{\min}(\omega_{\mathcal{A}}) > 0$ is equivalent to $\omega_{\mathcal{A}}$ being ample (see [Barton 1971]).

**Remark 2.3.** Theorem 2.1 holds more generally if $k$ is only supposed to be perfect (the proof does not use the fact that $k$ is finite).

**Remark 2.4.** It would be interesting to provide an explicit example of an abelian variety $A$ as in the introduction to this article, such that $A$ is ordinary, $\mathcal{A}$ is semiabelian, $\text{Tr}_{\bar{K}|\bar{k}}(A_{\bar{K}}) = 0$ and $G_{\mathcal{A}} \neq 0$. It should be possible to construct such an example by considering mod $p$ reductions of the abelian variety constructed in [Catanese and Dettweiler 2016, Theorem 1.3]. We hope to return to this question in a later article. The following question is also of interest: is there an ordinary abelian variety $A$ as above, such that $A$ has maximal Kodaira–Spencer rank, $\mathcal{A}$ is semiabelian and $G_{\mathcal{A}} \neq 0$?

**Proposition 2.5.** *Suppose that $A$ is ordinary and that $\mathcal{A}$ is semiabelian. Suppose that $A(K^{\text{perf}})$ is not finitely generated. Then $G_{\mathcal{A}}$ is of order $> 1$ and $\mathcal{A}/G_{\mathcal{A}}$ is also semiabelian.*

**Proposition 2.6.** *Suppose that $A$ is ordinary and that $\mathcal{A}$ is semiabelian over $S$. Suppose that $A(K^{\mathrm{perf}})$ is not finitely generated.*

*Then there is a finite flat morphism*

$$\phi : \mathcal{A} \to \mathcal{B},$$

*where $\mathcal{B}$ is a semiabelian over $S$ and a finite flat morphism*

$$\lambda : \mathcal{B} \to \mathcal{B}$$

*such that* $\ker(\phi)$ *and* $\ker(\lambda)$ *are multiplicative group schemes and such that the order of* $\ker(\lambda)$ *is* $> 1$.

### 2B. *Consequences of infiniteness of* $\mathbf{Tor}_p(A(K^{\mathbf{sep}}))$ *or* $\mathbf{Tor}_p(A(K^{\mathbf{unr}}))$.

**Theorem 2.7.** *Suppose that* $\mathrm{Tr}_{\overline{K}|\bar{k}}(A_{\overline{K}}) = 0$. *Suppose that the action of* $\mathrm{Gal}(K^{\mathrm{sep}}|K)$ *on* $\mathrm{Tor}_p(A(K^{\mathrm{unr}}))$ *factors through* $\mathrm{Gal}(K^{\mathrm{sep}}|K)^{\mathrm{ab}}$. *Then* $\mathrm{Tor}_p(A(K^{\mathrm{unr}}))$ *is finite.*

Here $\mathrm{Gal}(K^{\mathrm{sep}}|K)^{\mathrm{ab}}$ is the maximal abelian quotient of $\mathrm{Gal}(K^{\mathrm{sep}}|K)$.

**Proposition 2.8.** *Suppose that* $\dim(A) \leqslant 2$ *and that* $\mathrm{Tr}_{\overline{K}|\bar{k}}(A_{\overline{K}}) = 0$. *Then* $\mathrm{Tor}_p(A(K^{\mathrm{unr}}))$ *is finite.*

**Theorem 2.9.** *Suppose that* $\mathrm{Tor}_p(A(K^{\mathrm{sep}}))$ *is infinite. Then there is an étale $K$-isogeny*

$$\phi : A \to B,$$

*where $B$ is an abelian variety over $K$ and there is an étale $K$-isogeny*

$$\lambda : B \to B$$

*such that the order of* $\ker(\lambda)$ *is* $> 1$ *and such that the orders of* $\ker(\lambda)$ *and* $\ker(\phi)$ *are powers of $p$.*

**Theorem 2.10.** *Suppose that there exists an étale $K$-isogeny $\phi : A \to A$, such that* $\deg(\phi) = p^r$ *for some* $r > 0$. *Suppose also that $A$ is a geometrically simple abelian variety and that $\mathcal{A}$ is a semiabelian scheme.*

*Then $\mathcal{A}$ is an abelian scheme and $\phi$ extends to an étale (necessarily finite) $S$-morphism $\mathcal{A} \to \mathcal{A}$ of group schemes.*

## 3. Semistable sheaves on curves

Let $Y$ be a scheme, which is smooth, projective and geometrically connected of relative dimension one over a field $t_0$.

Suppose to begin with that $t_0$ is algebraically closed.

If $V$ is a nonzero coherent locally free sheaf on $Y$, we write (as is customary)

$$\mu(V) = \deg(V)/\,\mathrm{rk}(V),$$

where

$$\deg(V) := \int_Y \mathrm{c}_1(V)$$

and $\mathrm{rk}(V)$ is the rank of $V$. The quantity $\mu(V)$ is called the *slope* of $V$. Recall that a nonzero locally free coherent sheaf $V$ on $Y$ is called semistable if for any nonzero coherent subsheaf $W \subseteq V$, we have

$\mu(W) \leqslant \mu(V)$. Let $V/Y$ be a nonzero locally free coherent sheaf on $Y$. There is a unique filtration by coherent subsheaves

$$0 = V_0 \subsetneq V_1 \subsetneq V_2 \subsetneq \cdots \subsetneq V_{\mathrm{hn}(V)} = V$$

such that all the sheaves $V_i/V_{i-1}$ ($1 \leqslant i \leqslant \mathrm{hn}(V)$) are (locally free and) semistable and such that the sequence $\mu(V_i/V_{i-1})$ is strictly decreasing. This filtration is called the *Harder–Narasimhan filtration* of $V$ (shorthand: HN filtration). One then defines

$$V_{\min} := V/V_{\mathrm{hn}(V)-1}, \quad V_{\max}(V) := V_1 \qquad \text{and} \qquad \mu_{\max}(V) := \mu(V_1), \quad \mu_{\min}(V) := \mu(V_{\min}).$$

Let now $r \in \mathbb{Q}$. Suppose that $r \in \{\mu(V_1), \ldots, \mu(V/V_{\mathrm{hn}(V)-1})\}$. Let $i(r) \in \mathbb{N}$ be the unique natural number such that $\mu(V_{i(r)}/V_{i(r)-1}) = r$. We shall write

$$V_{=r} := V_{i(r)}/V_{i(r)-1} \qquad \text{and} \qquad V_{\geq r} := V_{i(r)}.$$

We shall also write

$$V_{>r} := V_{j(r)},$$

where $j(r) \in \mathbb{N}$ is the largest natural number such that $\mu(V_{j(r)}/V_{j(r)-1}) > r$.

One basic property of semistable sheaves that we shall use repeatedly is the following. If $V$ and $W$ are nonzero coherent locally free sheaves on $Y$ and $\mu_{\min}(V) > \mu_{\max}(W)$ then $\mathrm{Hom}_Y(V, W) = 0$. This follows from the definitions.

See [Brenner et al. 2008, Chapter 5] (for instance) for all these notions.

If $V$ is a nonzero coherent locally free sheaf on $Y$ and $t_0$ has positive characteristic, we say that $V$ is *Frobenius semistable* if $F_Y^{\circ r,*}(V)$ is semistable for all $r \in \mathbb{N}$. The terminology *strongly semistable* also appears in the literature.

**Theorem 3.1.** *Let $V$ be a nonzero coherent locally free sheaf on $Y$. There is an $\ell_0 = \ell_0(V) \in \mathbb{N}$ such that the quotients of the Harder–Narasimhan filtration of $F_Y^{\circ \ell_0,*}(V)$ are all Frobenius semistable.*

*Proof.* See, e.g., [Langer 2004, Theorem 2.7, p. 259]. □

Theorem 3.1 shows in particular that the definitions

$$\bar{\mu}_{\min}(V) := \lim_{\ell\to\infty} \mu_{\min}(F_Y^{\circ \ell,*}(V))/p^\ell, \quad \bar{\mu}_{\max}(V) := \lim_{\ell\to\infty} \mu_{\max}(F_Y^{\circ \ell,*}(V))/p^\ell,$$
$$\overline{\mathrm{rk}}_{\min}(V) := \lim_{\ell\to\infty} \mathrm{rk}((F_Y^{\circ \ell,*}(V))_{\min}), \quad \overline{\mathrm{rk}}_{\max}(V) := \lim_{\ell\to\infty} \mathrm{rk}((F_Y^{\circ \ell,*}(V))_{\max}),$$

make sense if $V$ is a nonzero locally free and coherent sheaf on $Y$.

*Suppose now that $t_0$ is only perfect* (not necessarily algebraically closed). If $V$ is a nonzero coherent sheaf on $Y$, then we shall write $\mu(V) := \mu(V_{\bar{t}_0})$ and we shall say that $V$ is semistable if $V_{\bar{t}_0}$ is semistable. The HN filtration of $V_{\bar{t}_0}$ is invariant under $\mathrm{Gal}(\bar{t}_0|t_0)$ by unicity and by a simple descent argument, we see that there is a unique filtration by coherent subsheaves

$$V_0 \subsetneq V_1 \subsetneq V_2 \subsetneq \cdots \subsetneq V_{\mathrm{hn}(V)}$$

such that

$$V_{0,\bar{t}_0} \subsetneq V_{1,\bar{t}_0} \subsetneq V_{2,\bar{t}_0} \subsetneq \cdots \subsetneq V_{\mathrm{hn}(V),\bar{t}_0}$$

is the HN filtration of $V_{\bar{t}_0}$. We then define as before

$$\mu_{\max}(V) := \mu(V_1) \qquad \text{and} \qquad \mu_{\min}(V) := \mu(V/V_{\mathrm{hn}(V)-1}).$$

Notice that we have $\mu_{\max}(V) = \mu_{\max}(V_{\bar{t}_0})$ and $\mu_{\min}(V) = \mu_{\min}(V_{\bar{t}_0})$.

Notice that if $V$ and $W$ are nonzero coherent locally free coherent sheaves on $Y$ and $\mu_{\min}(V) > \mu_{\max}(W)$ then we still have $\mathrm{Hom}_Y(V, W) = 0$, since there is a natural inclusion

$$\mathrm{Hom}_Y(V, W) \subseteq \mathrm{Hom}_{Y_{\bar{t}_0}}(V_{\bar{t}_0}, W_{\bar{t}_0}).$$

If $t_0$ has positive characteristic, we shall say that $V$ is Frobenius semistable if $V_{\bar{t}_0}$ is Frobenius semistable. Since Frobenius morphisms commute with all morphisms, this is equivalent to requiring that $F_Y^{r,*}(V)$ is semistable for all $r \in \mathbb{N}$ (with our extended definition of semistability).

We can now extend the range of the terminology introduced above:

$$\begin{aligned}
V_{\min} &:= V/V_{\mathrm{hn}(V)-1}, & V_{\max} &:= V_1,\\
\bar{\mu}_{\min}(V) &:= \lim_{\ell\to\infty} \mu_{\min}(F_Y^{\circ\ell,*}(V))/p^\ell, & \bar{\mu}_{\max}(V) &:= \lim_{\ell\to\infty} \mu_{\max}(F_Y^{\circ\ell,*}(V))/p^\ell,\\
\overline{\mathrm{rk}}_{\min}(V) &:= \lim_{\ell\to\infty} \mathrm{rk}((F_Y^{\circ\ell,*}(V))_{\min}), & \overline{\mathrm{rk}}_{\max}(V) &:= \lim_{\ell\to\infty} \mathrm{rk}((F_Y^{\circ\ell,*}(V))_{\max}).
\end{aligned}$$

Note that we have $\bar{\mu}_{\min}(V) = \bar{\mu}_{\min}(V_{\bar{t}_0})$, $\bar{\mu}_{\max}(V) = \bar{\mu}_{\max}(V_{\bar{t}_0})$, $\overline{\mathrm{rk}}_{\min}(V) = \overline{\mathrm{rk}}_{\min}(V_{\bar{t}_0})$, $\overline{\mathrm{rk}}_{\max}(V) = \overline{\mathrm{rk}}_{\max}(V_{\bar{t}_0})$ as expected.

If $V$ is a nonzero coherent locally free coherent sheaf on $Y$ such that all the quotients of the HN filtration of $V$ are Frobenius semistable, we shall say that $V$ has a Frobenius semistable HN filtration. Note that by Theorem 3.1 above, for any nonzero coherent locally free coherent sheaf $V$ on $Y$, the sheaf $F^{\circ r,*}(V)$ has a Frobenius semistable HN filtration for all but finitely many $r \in \mathbb{N}$.

The following simple lemma will also prove very useful. It was suggested by J.-B. Bost.

**Lemma 3.2.** *Let $V$ and $W$ be coherent locally free sheaves on $Y$. Suppose that $\mu(V) = \mu(W)$ and that $\mathrm{rk}(V) = \mathrm{rk}(W)$. Let $\phi : V \to W$ be a monomorphism of $\mathcal{O}_Y$-modules. Then $\phi$ is an isomorphism.*

*Proof.* We may suppose that $V$ and $W$ are of positive rank, otherwise the lemma is tautologically true. Let $M := \det(W) \otimes \det(V)^\vee$. The assumptions imply that $\deg(M) = 0$. Let $\det(\phi) \in H^0(Y, M)$ be the section induced by $\phi$. The zero scheme $Z(\det(\phi))$ of $\det(\phi)$ is a torsion sheaf since $\det(\phi)$ is nonzero at the generic point of $Y$ and the length of $Z(\det(\phi))$ is equal to the degree of $M$ so $Z(\det(\phi))$ must be empty. In other words, $M$ is the trivial sheaf and $\det(\phi)$ is a constant nonzero section of $M$. In particular, $\phi$ is an isomorphism. □

## 4. Finite flat group schemes over curves

The terminology of this section is independent of the introduction.

**4A.** ***Quotients by proper flat group schemes.*** Let $Y$ be a noetherian scheme. Let $G$ be a commutative strongly quasiprojective flat group scheme over $Y$. See [Bosch et al. 1990, 8.2, p. 211] for the definition of strong quasiprojectivity. Note that if $Y$ is regular then $G$ is strongly quasiprojective over $Y$ if it is quasiprojective over $Y$.

Suppose that $H$ is a closed subgroup scheme of $G$, which is proper and flat over $Y$. The $Y$-scheme $G$ (resp. $H$) defines a functor $\underline{G}$ (resp. $\underline{H}$) from the category of $Y$-schemes to the category of abelian groups. Both functors are fppf sheaves by a classical result of Grothendieck. We may thus form the quotient $\underline{G}/\underline{H}$ of $\underline{G}$ and $\underline{H}$ in the category of fppf sheaves.

The following proposition describes the quotient construction that we use in this text.

**Proposition 4.1.** *The fppf sheaf $\underline{G}/\underline{H}$ is representable by a group scheme $G/H$ over $Y$, which is also strongly quasiprojective. The natural morphism $q : G \to G/H$ is proper and faithfully flat and makes $G$ into an $H_{G/H}$-torsor over $G/H$.*

*Proof.* See [Bosch et al. 1990, Theorem 8.12, p. 220]. □

Note that if $G$ is semiabelian and $Y$ is normal then $G$ is quasiprojective over $Y$ (combine [Moret-Bailly 1985, VI.3.1] with [Raynaud 1970, XI.1.4]). In particular if $Y$ is regular and $G$ is semiabelian then $G$ is strongly quasiprojective over $Y$.

**4B.** ***The* HN *-filtration on the Lie algebra of a finite flat group scheme of height one.*** Let $S$ be a smooth, projective and geometrically connected curve over a perfect field $k$. Suppose that $\mathrm{char}(k) = p > 0$.

The following preliminary lemma will be very useful.

**Lemma 4.2.** *Let $G$ be a finite flat commutative group scheme over $S$. Let $T \to S$ be a flat, radicial and finite morphism and let $\phi : H \hookrightarrow G_T$ be a closed subgroup scheme, which is finite, flat and multiplicative. Then there is a finite flat closed subgroup scheme $\phi_0 : H_0 \hookrightarrow G$, such that $\phi_{0,T} \simeq \phi$.*

*Proof.* Taking Cartier duals, we get a morphism

$$\phi^\vee : G_T^\vee \to H^\vee.$$

Notice that $H^\vee$ is étale over $T$, since $H$ is multiplicative. By radicial invariance of étale morphisms, there is a finite flat group scheme $J_0 \to S$, such that $J_{0,T} \simeq H^\vee$. Notice also that the morphism $\phi^\vee$ is given by a section of the first projection

$$G_T^\vee \times_T H^\vee \to G_T^\vee$$

and since $H^\vee$ is étale over $T$, the image of this section is open and closed (see [Milne 1980, Corollary 3.12]). Since the projection morphism

$$G_T^\vee \times_T H^\vee \to G^\vee \times_S J_0$$

is also radicial, this open set comes from a unique open subset of $G \times_S J_0$ and this open subset defines an open and closed subscheme of $G^\vee \times_S J_0$, which is isomorphic to $G^\vee$ via the first projection. Hence the morphism $\phi^\vee$ comes from a unique morphism $G^\vee \to J_0$. Taking the Cartier dual of this morphism gives the morphism $\phi_0$. □

Recall that a commutative finite flat group scheme $\psi : G \to S$ over $S$ is said to be *of height one* if $F_{G/S} = \epsilon_{G/S} \circ \psi$. Recall also that a (sheaf in) commutative $p$-Lie (resp. $p$-coLie) algebras $V$ over $S$ is a coherent locally free sheaf $V$ on $S$ together with a morphism of $\mathcal{O}_S$-modules $F_S^*(V) \to V$ (resp. $V \to F_S^*(V)$). A morphism of commutative $p$-Lie (resp. $p$-coLie) algebras $V \to W$ is a morphism of $\mathcal{O}_S$-modules from $V$ to $W$ satisfying an evident compatibility condition. There is a covariant functor $\mathrm{Lie}(\,\cdot\,)$ (resp. contravariant functor $\mathrm{coLie}(\,\cdot\,)$) from the category of commutative finite flat group schemes of height one over $S$ to the category of commutative $p$-Lie (resp. $p$-coLie) algebras, which sends a group scheme $G$ over $S$ to $\mathrm{Lie}(G) := \epsilon_{G/S}^*(\Omega_{G/S})^\vee$ (resp. $\mathrm{coLie}(G) := \epsilon_{G/S}^*(\Omega_{G/S})$, together with the morphism

$$\mathrm{Lie}(V_{G^{(p)}/S}) := (V^*_{G^{(p)}/S})^\vee : F_S^*(\mathrm{Lie}(G)) = \mathrm{Lie}(G^{(p)}) \to \mathrm{Lie}(G)$$
$$\big(\text{resp. } \mathrm{coLie}(V_{G^{(p)}/S}) := V^*_{G^{(p)}/S} : \mathrm{coLie}(G) \to F_S^*(\mathrm{coLie}(G^{(p)})) = \mathrm{coLie}(G^{(p)})\big).$$

Here $(V^*_{G^{(p)}/S})^\vee$ (resp. $V^*_{G^{(p)}/S}$) is the dual of the pull-back morphism $V^*_{G^{(p)}/S}$ (resp. is the pull-back morphism) on differentials induced by the Verschiebung morphism $V_{G^{(p)}/S}$.

The category of sheaves in commutative $p$-Lie algebras is tautologically antiequivalent to the category of sheaves in commutative $p$-coLie algebras.

It can be shown that Lie is an equivalence of additive categories (see [SGA $3_1$ 2011, Exposé VIIA, Remark 7.5]). In particular, a sequence of finite flat group schemes of height one

$$0 \to G' \to G \to G'' \to 0$$

is exact if and only if the sequence

$$0 \to \mathrm{Lie}(G') \to \mathrm{Lie}(G) \to \mathrm{Lie}(G'') \to 0$$

is a sequence of commutative $p$-Lie algebras. Furthermore, we have

$$\mathrm{order}(G) = p^{\mathrm{rk}(\mathrm{Lie}(G))}$$

(see [Mumford 1970, Proof of Theorem, p. 139, paragraph 14]).

**Lemma 4.3.** *Let $\phi : V \to W$ be a morphism of commutative $p$-Lie algebras. Then the image* $\mathrm{Im}(\phi)$ *(resp. the kernel* $\ker(\phi)$*) of $\phi$ as a morphism of $\mathcal{O}_S$-modules is endowed with a unique structure of commutative $p$-Lie algebra, such that the morphism* $\mathrm{Im}(\phi) \to W$ *(resp.* $\ker(\phi) \to V$*) is a morphism of commutative $p$-Lie algebras.*

*Proof.* The proof is left to the reader. □

If $\phi : V \to W$ is an injective morphism of commutative $p$-Lie algebras, we shall say that $\mathrm{Im}(\phi)$ is a subsheaf in commutative $p$-Lie algebras. Beware that in this situation, the arrow $\phi$ might have no cokernel in the category of commutative $p$-Lie algebras. So in particular, $\mathrm{Im}(\phi)$ might not correspond to a subgroup scheme. On the other hand, if the quotient of $\mathcal{O}_S$-modules $W/\mathrm{Im}(\phi)$ is locally free, then $W/\mathrm{Im}(\phi)$ can be endowed with an evident commutative $p$-Lie algebra structure, making it into a cokernel of $W$ by $\mathrm{Im}(\phi)$ in the category of commutative $p$-Lie algebras. In that case, $\mathrm{Im}(\phi)$ corresponds to a subgroup scheme.

We shall say that a finite flat commutative group scheme $G$ of height one (or its associated commutative $p$-Lie algebra) is *biinfinitesimal* if the associated morphism $F_S^*(\mathrm{Lie}(G)) \to \mathrm{Lie}(G)$ is nilpotent. To say that $F_S^*(\mathrm{Lie}(G)) \to \mathrm{Lie}(G)$ is nilpotent means that for some $n \geq 1$, the composition

$$F_S^{\circ n,*}(\mathrm{Lie}(G)) \to F_S^{\circ(n-1),*}(\mathrm{Lie}(G)) \to \cdots \to F_S^*(\mathrm{Lie}(G)) \to \mathrm{Lie}(G) \to 0$$

vanishes. We notice without proof that if

$$0 \to G' \to G \to G'' \to 0$$

is an exact sequence of commutative finite flat group schemes, then $G'$ and $G''$ are biinfinitesimal if and only if $G$ is biinfinitesimal. Note also that a finite flat commutative group scheme $G$ of height one is multiplicative if and only if the associated morphism $F_S^*(\mathrm{Lie}(G)) \to \mathrm{Lie}(G)$ is an isomorphism. This implies that if $G_1$ and $G_2$ are finite flat group schemes of height one over $S$, where $G_1$ is biinfinitesimal and $G_2$ is multiplicative then there are no nonzero morphisms of group schemes from $G_1$ to $G_2$ and also no nonzero morphisms of group schemes from $G_2$ to $G_1$.

We inserted the following alternative proof of a special case of Lemma 4.2 to show the mechanics of $p$-Lie algebras at work in a simple situation.

*Second proof of Lemma 4.2 when $G$ is of height one and $T$ is smooth.* We may assume that $T \simeq S$ and that $T \to S$ is a power $F_S^{\circ n}$ of $F_S$. By induction on $n$, we are reduced to prove the statement for $n = 1$.

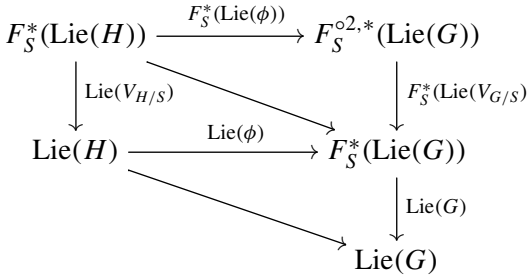
We are given a commutative diagram with exact rows and columns

$$\begin{array}{ccccc}
 & & 0 & & \\
 & & \downarrow & & \\
0 & \longrightarrow & F_T^*(\mathrm{Lie}(H)) & \xrightarrow{F_T^*(\mathrm{Lie}(\phi))} & F_T^*(\mathrm{Lie}(G)_T) \\
 & & \downarrow{\scriptstyle \mathrm{Lie}(V_{H/T})} & & \downarrow{\scriptstyle \mathrm{Lie}(V_{G_T/T})} \\
0 & \longrightarrow & \mathrm{Lie}(H) & \xrightarrow{\mathrm{Lie}(\phi)} & \mathrm{Lie}(G)_T \\
 & & \downarrow & & \\
 & & 0 & &
\end{array}$$

With the above reductions in place, this gives a commutative diagram with exact rows and columns

$$\begin{array}{ccccc}
 & & 0 & & \\
 & & \downarrow & & \\
0 & \longrightarrow & F_S^*(\mathrm{Lie}(H)) & \xrightarrow{F_S^*(\mathrm{Lie}(\phi))} & F_S^{\circ 2,*}(\mathrm{Lie}(G)) \\
 & & \downarrow{\scriptstyle \mathrm{Lie}(V_{H/S})} & & \downarrow{\scriptstyle F_S^*(\mathrm{Lie}(V_{G/S})} \\
0 & \longrightarrow & \mathrm{Lie}(H) & \xrightarrow{\mathrm{Lie}(\phi)} & F_S^*(\mathrm{Lie}(G)) \\
 & & \downarrow & & \\
 & & 0 & &
\end{array}$$

Now consider the commutative diagram




where the diagonal arrows are defined so that the diagram becomes commutative. The labelling of the arrows shows that the upper triangle is the base change by $F_S$ of the lower triangle. Hence the image of $\mathrm{Lie}(\phi)$ is the base change by $F_S$ of the image of $\mathrm{Lie}(H)$ in $\mathrm{Lie}(G)$, since $\mathrm{Lie}(V_{H/S})$ is an isomorphism. So $H_0$ can be defined as the group scheme of height one associated with the image of $\mathrm{Lie}(H)$ in $\mathrm{Lie}(G)$. □

**Lemma 4.4.** *Let $V$ be a sheaf in commutative $p$-Lie algebras $V$ over $S$. Suppose that the* HN*filtration*

$$0 = V_0 \subsetneq V_1 \subsetneq V_2 \subsetneq \cdots \subsetneq V_{\mathrm{hn}(V)} = V$$

*of $V$ is Frobenius semistable. Then for any $V_i$ such that $\mu_{\min}(V_i) \geq 0$, $V_i$ is a subsheaf in commutative $p$-Lie algebras $V$ over $S$. If $\mu_{\min}(V_i) > 0$ then $V_i$ is biinfinitesimal.*

*Proof.* For the first statement, consider the morphism $\phi : F_S^*(V_i) \to V$ given by the composition of the inclusion $F_S^*(V_i) \to F_S^*(V)$ with the morphism $F_S^*(V) \to V$ given by the commutative $p$-Lie algebra structure. We have to check that the image of $\phi$ lies in $V_i$. The composition of $\phi$ with the quotient morphism $V \to V/V_i$ gives a morphism $F_S^*(V_i) \to V/V_i$ and it is equivalent to check that this morphism vanishes. Now we compute

$$\mu_{\min}(F_S^*(V_i)) = p \cdot \mu(V_i/V_{i-1}) \quad \text{and} \quad \mu_{\max}(V/V_i) = \mu(V_{i+1}/V_i) < \mu(V_i/V_{i-1}),$$

and thus $\mu_{\min}(F_S^*(V_i)) > \mu_{\max}(V/V_i)$. We conclude that $\mathrm{Hom}_S(F_S^*(V_i), V/V_i) = 0$ (see the discussion after Theorem 3.1) which concludes the proof of the first statement. To prove the second statement, it is sufficient by the remarks preceding the lemma to show that $V_i/V_{i-1}$ is biinfinitesimal for all indices $i$ such that $\mu(V_i/V_{i-1}) > 0$. By the above computation, we have

$$\mu_{\min}(F_S^*(V_i/V_{i-1})) = \mu(F_S^*(V_i/V_{i-1})) = p \cdot \mu(V_i/V_{i-1})$$

and thus $\mu_{\min}(F_S^*(V_i/V_{i-1})) > \mu(V_i/V_{i-1})$. Again, this implies that $\mathrm{Hom}_S(F_S^*(V_i/V_{i-1}), V_i/V_{i-1}) = 0$, showing that $V_i/V_{i-1}$ is biinfinitesimal. □

**Remark 4.5.** As explained in the introduction, a characteristic 0 analogue of Lemma 4.4 can be found in [Bost 2004, Lemma 2.9]. See also [Shepherd-Barron 1992, Lemma 9.1.3.1], where a variant of a special case of Lemma 4.4 is proven under the assumption that $p$ is sufficiently large.

**Lemma 4.6.** *Let $G$ be a commutative finite flat group scheme of height one over $S$ and suppose given an exact sequence*

$$0 \to G_{\mathrm{binf}} \to G \to G_{\mu} \to 0$$

*of finite flat group schemes such that $G_{\mu}$ is multiplicative and $G_{\mathrm{binf}}$ is biinfinitesimal. Then the sequence splits and this splitting is unique.*

*Proof.* Consider the commutative diagram with exact rows and columns

$$\begin{array}{ccccccccc}
0 & \longrightarrow & \ker(\mathrm{Lie}(V^{(n)}_{G_{\mathrm{binf}}{}^{(p^n)}/S})) & \longrightarrow & \ker(\mathrm{Lie}(V^{(n)}_{G^{(p^n)}/S})) & \longrightarrow & 0 & & \\
 & & \downarrow \simeq & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & F_S^{\circ n,*}(\mathrm{Lie}(G_{\mathrm{binf}})) & \longrightarrow & F_S^{\circ n,*}(\mathrm{Lie}(G)) & \longrightarrow & F_S^{\circ n,*}(\mathrm{Lie}(G_{\mu})) & \longrightarrow & 0 \\
 & & \downarrow =0 & & \downarrow & & \downarrow \simeq & & \\
0 & \longrightarrow & \mathrm{Lie}(G_{\mathrm{binf}}) & \longrightarrow & \mathrm{Lie}(G) & \longrightarrow & \mathrm{Lie}(G_{\mu}) & \longrightarrow & 0
\end{array}$$

where $n \geq 0$ is chosen so that $V^{(n),*}_{G_{\mathrm{binf}}{}^{(p^n)}/S} = 0$. Then the image of the arrow

$$F_S^{\circ n,*}(\mathrm{Lie}(G)) \to \mathrm{Lie}(G)$$

splits the bottom sequence. For the unicity of the splitting, note that for any two splittings $\sigma_1, \sigma_2$ of the bottom sequence the morphism $\sigma_1 - \sigma_2 : \mathrm{Lie}(G_{\mu}) \to \mathrm{Lie}(G)$ of vector bundles factors through the image of $\mathrm{Lie}(G_{\mathrm{binf}})$. It thus defines a morphism of vector bundles $\mathrm{Lie}(G_{\mu}) \to \mathrm{Lie}(G_{\mathrm{binf}})$, which is by construction a morphism of $p$-Lie algebras. Such a morphism must vanish (see the discussion after Lemma 4.3). Thus $\sigma_1 = \sigma_2$. □

**Lemma 4.7.** *Let $G$ be a commutative finite flat group scheme of height one over $S$. Suppose that* $\mathrm{Lie}(G)$ *is Frobenius semistable of slope* $0$*. Let $n \geq 0$ be such that* $\mathrm{rk}(\ker(V^{(n),*}_{G^{(p^n)}/S}))$ *is maximal. Then there is a canonical decomposition*

$$G^{(p^n)} \simeq H_{\mathrm{binf}} \times_S H_{\mu},$$

*where $H_{\mathrm{binf}}$ (resp. $H_{\mu}$) is a biinfinitesimal* (*resp. multiplicative*) *finite flat group scheme over $S$.*

*Proof.* Consider the commutative diagram

$$\begin{array}{ccccccccc}
0 & \longrightarrow & F_S^{\circ n,*}(\ker(\mathrm{Lie}(V^{(n)}_{G^{(p^n)}/S}))) & \longrightarrow & F_S^{\circ n,*}(\ker(\mathrm{Lie}(V^{(n)}_{G^{(p^n)}/S}))) & \longrightarrow & 0 & & \\
 & & \downarrow \sim & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & F_S^{\circ n,*}(\ker(\mathrm{Lie}(V^{(n)}_{G^{(p^n)}/S}))) & \longrightarrow & F_S^{\circ (2n),*}(\mathrm{Lie}(G)) & \longrightarrow & F_S^{\circ n,*}(W) & \longrightarrow & 0 \\
 & & \downarrow =0 & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & \ker(\mathrm{Lie}(V^{(n)}_{G^{(p^n)}/S})) & \longrightarrow & F_S^{\circ n,*}(\mathrm{Lie}(G)) & \longrightarrow & W & \longrightarrow & 0
\end{array}$$

where $n \geq 0$ is such that $\mathrm{rk}(\ker(\mathrm{Lie}(V^{(n)}_{G^{(p^n)}/S})))$ is maximal and $W$ is the image of $\mathrm{Lie}(V^{(n)}_{G^{(p^n)}/S})$. The two bottom rows and the two leftmost columns in this diagram are exact by construction. Furthermore the map $F_S^{(n),*}W \to W$ is a monomorphism for otherwise $\mathrm{rk}(\ker(\mathrm{Lie}(V^{(n)}_{G^{(p^n)}/S})))$ is not maximal. The diagram thus has exact rows and columns. Since the second row gives a surjection

$$F_S^{\circ(2n),*}(\mathrm{Lie}(G)) \to F_S^{\circ n,*}(W)$$

we have $\mu_{\min}(F_S^{\circ n,*}(W)) \geq 0$. Also, since the second column gives an injection

$$F_S^{\circ n,*}(W) \hookrightarrow F_S^{(n),*}(\mathrm{Lie}(G))$$

we have $\mu_{\max}(F_S^{\circ n,*}(W)) \leq 0$. Thus $F_S^{\circ n,*}(W)$ is of slope 0. Thus $W$ is also of slope 0. Hence by Lemma 3.2, the monomorphism

$$F_S^{\circ n,*}(W) \to W$$

is an isomorphism. Now we see that the image of the morphism $F_S^{\circ(2n),*}(\mathrm{Lie}(G)) \to F_S^{\circ n,*}(\mathrm{Lie}(G))$ splits the bottom sequence. □

**Lemma 4.8.** *Let $G$ be a finite flat commutative group scheme of height one over $S$. There exists a* (*necessarily unique*) *multiplicative subgroup scheme $G_\mu \hookrightarrow G$, such that if $H$ is a multiplicative subgroup scheme of height one over $S$ and $f : H \to G$ is a morphism of group schemes, then $f$ factors through $G_\mu$. Furthermore, for any $n \geq 0$, we have $(G_\mu)^{(p^n)} = (G^{(p^n)})_\mu$. If $G$ is multiplicative over a dense open subset of $S$ and* $\mathrm{Lie}(G)$ *has Frobenius semistable* HN *filtration then* $\mathrm{Lie}(G) = \mathrm{Lie}(G)_{\leq 0}$ *and $G_\mu$ corresponds to the subgroup scheme associated with* $\mathrm{Lie}(G)_{=0}$.

*Proof.* In view of Lemma 4.2, we may replace $G$ by $G^{(p^n)}$ for any $n \geq 0$ and in particular suppose that $\mathrm{Lie}(G)$ has a Frobenius semistable HN filtration. Let $f : H \to G$ be a morphism of group schemes and consider the corresponding map

$$\mathrm{Lie}(f) : \mathrm{Lie}(H) \to \mathrm{Lie}(G).$$

Since $H$ is multiplicative, $\mathrm{Lie}(H)$ is Frobenius semistable of slope 0 (this is a consequence of Theorem 3.1). Thus the image of $\mathrm{Lie}(f)$ lies in $\mathrm{Lie}(G)_{\geq 0}$. According to Lemma 4.4 there is an exact sequence of $p$-Lie algebras

$$0 \to \mathrm{Lie}(G)_{>0} \to \mathrm{Lie}(G)_{\geq 0} \xrightarrow{\pi} \mathrm{Lie}(G)_{=0} \to 0$$

and we may assume according to Lemma 4.7 that there is a splitting

$$\mathrm{Lie}(G)_{=0} \simeq \mathrm{Lie}(G)_{=0,\mathrm{binf}} \oplus \mathrm{Lie}(G)_{=0,\mu}$$

of $\mathrm{Lie}(G)_{=0}$ into multiplicative and biinfinitesimal part (we might have to twist $G$ some more for this). The inverse image of $\mathrm{Lie}(G)_{=0,\mu}$ by $\pi$ gives a $p$-Lie subalgebra $\pi^*(\mathrm{Lie}(G)_{=0,\mu})$ of $\mathrm{Lie}(G)_{\geq 0}$. This gives an exact sequence

$$0 \to \pi^*(\mathrm{Lie}(G)_{=0,\mu}) \to \mathrm{Lie}(G)_{\geq 0} \to \mathrm{Lie}(G)_{=0,\mathrm{binf}} \to 0$$

Since $\mathrm{Lie}(H)$ is multiplicative, the image of $\mathrm{Lie}(H)$ in $\mathrm{Lie}(G)_{=0,\mathrm{binf}}$ vanishes and thus the image of $\mathrm{Lie}(H)$ lies in $\pi^*(\mathrm{Lie}(G)_{=0,\mu})$. On the other hand by Lemma 4.6 and Lemma 4.4, we have again a canonical decomposition

$$\pi^*(\mathrm{Lie}(G)_{=0,\mu})_\mu \oplus \pi^*(\mathrm{Lie}(G)_{=0,\mu})_{\mathrm{binf}}$$

into multiplicative and biinfinitesimal part and thus the image of $\mathrm{Lie}(f)$ lies in $\pi^*(\mathrm{Lie}(G)_{=0,\mu})_\mu$. Now $\pi^*(\mathrm{Lie}(G)_{=0,\mu})_\mu$ is a multiplicative $p$-Lie subalgebra of $\mathrm{Lie}(G)$ and it defines the required subgroup scheme.

If $G$ is multiplicative over an open subset of $S$ then we have an injection

$$F_S^{on,*}(\mathrm{Lie}(G)) \hookrightarrow \mathrm{Lie}(G)$$

(obtained by composition) for any $n \geq 0$ and thus if $\mathrm{Lie}(G)$ has Frobenius semistable HN filtration then we must have $\mathrm{Lie}(G) = \mathrm{Lie}(G)_{\leq 0}$. Secondly the morphism $F_S^*(\mathrm{Lie}(G)) \hookrightarrow \mathrm{Lie}(G)$ then induces an injection

$$F_S^*(\mathrm{Lie}(G)_{=0}) \hookrightarrow \mathrm{Lie}(G)_{=0}$$

and since both source and target in this map have the same rank and the same slope, we deduce from Lemma 3.2 that this map must be an isomorphism. Thus $\mathrm{Lie}(G)_{=0}$ is multiplicative and by the explicit construction above, it is associated with $G_\mu$. □

**Remark 4.9.** Note that the "connected étale" decomposition of $G_K^\vee$ (see the beginning of [Tate 1997]) gives a canonical exact sequence of group schemes

$$0 \to (G_K^\vee)_{\mathrm{inf}} \to G_K^\vee \to (G_K^\vee)_{\mathrm{et}} \to 0$$

over $K$, where $(G_K^\vee)_{\mathrm{inf}}$ is an infinitesimal group scheme and $(G_K^\vee)_{\mathrm{et}}$ is an étale group scheme over $K$. The group scheme $(G_K^\vee)_{\mathrm{et}}$ corresponds to a representation of $\mathrm{Gal}(K^{\mathrm{sep}}|K)$ into a finite $p$-group $E$ and one might be tempted to think that $G_\mu$ is the Cartier dual of the group scheme corresponding to the largest unramified quotient of $E$, i.e., the largest quotient of $E$, such that the action of $\mathrm{Gal}(K^{\mathrm{sep}}|K)$ factors through the fundamental group $\pi_1(S)$. This not so, however. Indeed, consider a finite flat commutative group scheme $G$ of height one, which is such that $\bar{\mu}_{\max}(\mathrm{Lie}(G)) < 0$. Then $G_\mu = 0$ and for any finite flat base change $S' \to S$, we also have $(G_{S'})_\mu = 0$. On the other hand $(G_K^\vee)_{\mathrm{et}}$ will become constant (and hence entirely unramified) after a finite separable field extension $K'|K$.

**4C.** ***Quotients of semiabelian schemes by finite flat multiplicative group schemes.*** Let $S$ be a smooth, projective and geometrically connected curve over a perfect field $k$.

**Lemma 4.10.** *Let $\mathcal{A} \to S$ be a semiabelian scheme. Suppose that there is an open dense subset $U \subseteq S$, such that $\mathcal{A}_U \to U$ is an abelian scheme. Suppose that $G \hookrightarrow \mathcal{A}$ is a finite, flat, closed subgroup scheme. Then the quotient scheme $\mathcal{A}/G$ is also a semiabelian scheme and $(\mathcal{A}/G)_U \to U$ is an abelian scheme.*

*Proof.* Since the quotient morphism $q : \mathcal{A} \to \mathcal{A}/G$ is faithfully flat, the group scheme $\mathcal{A}/G$ also has geometrically regular fibres (and is flat). Hence $\mathcal{A}/G$ is smooth over $S$. Over $U$, its fibres are proper since

the quotient morphism is also proper and they are thus abelian varieties. In other words, $(\mathcal{A}/G)_U \to U$ is an abelian scheme. Now let $s \in S$. Since $(\mathcal{A}/G)_s$ is smooth, we know by the Barsotti–Chevalley theorem (see [Milne 2017, Theorem 10.25, p. 157]) that $(\mathcal{A}/G)_s$ sits in the middle of an exact sequence

$$0 \to E_1 \to (\mathcal{A}/G)_s \to A_1 \to 0, \tag{1}$$

where $A_1$ is an abelian variety over $s$ and $E_1$ is a connected affine algebraic group variety over $s$. The subgroup variety $E_1$ is maximal among connected affine subgroup varieties of $(\mathcal{A}/G)_s$ (see [Milne 2017, Theorem 10.5, p. 153 and proof, and Theorem 10.24, p. 156]). Finally it has the form $E_1 = T_1 \times_s U$, where $T_1$ is a torus and $U$ is a connected unipotent group variety (see [Milne 2017, Chapter 10 (i), p. 161]). When we write that the sequence (1) is exact, we mean that the third morphism is faithfully flat and that its kernel is $E_1$.

By assumption, the corresponding presentation for $\mathcal{A}_s$ is

$$0 \to T \to \mathcal{A}_s \to A_0 \to 0,$$

where $T$ is a torus and $A_0$ is an abelian variety, both over $s$.

Let $D$ be the identity component of the closed subgroup scheme $q_s^{-1}(U \times 0)$ of $\mathcal{A}_s$ (see [Milne 2017, Proposition 1.14] for details). Since $s$ is perfect the closed subscheme $D_{\mathrm{red}}$ of $D$ is a closed subgroup scheme of $D$ (see [Milne 2017, Corollary 1.25, p. 24]). Moreover $D$ and hence $D_{\mathrm{red}}$ is affine, since $q_s$ is finite. Since $T$ is the maximal connected affine subgroup variety of $\mathcal{A}_s$, we see that $D_{\mathrm{red}}$ must be contained in $T$. However, every closed subgroup scheme of a multiplicative group over $s$ is multiplicative (see [SGA $3_{\mathrm{II}}$ 1970, 8.1, Exposé IX]) and thus $D_{\mathrm{red}}$ is multiplicative. Thus $D_{\mathrm{red}}$ is contained in the kernel of the morphism $q_s^{-1}(U \times 0) \to U \times 0$ (because there are no nontrivial morphisms between multiplicative and unipotent algebraic groups — see [Milne 2017, Corollary 15.18, p. 255]). Now notice that $q_s^{-1}(U \times 0)(\bar{s})/D(\bar{s})$ is a finite set (see [Milne 2017, Proposition 1.14, p. 21]). On the other hand $q_s(D(\bar{s})) = \{0\}$ by the above so $U(\bar{s})$ must be finite. Since $U$ is smooth, it must thus be trivial. This shows that $(\mathcal{A}/G)_s$ is an extension of an abelian variety by a torus. Since $s \in S$ was arbitrary, we see that $\mathcal{A}/G$ is a semiabelian scheme. □

**Lemma 4.11.** *Let $G \to S$ be a finite flat group scheme of multiplicative type. Then there is a finite étale morphism $T \to S$ such that $G_T$ is a diagonalisable group scheme.*

*Proof.* See [SGA $3_{\mathrm{II}}$ 1970, Exposé IX, Introduction]. □

**Lemma 4.12.** *Let $\mathcal{A} \to S$ be a smooth commutative group scheme. Suppose that $G \hookrightarrow \mathcal{A}$ is a finite, flat, closed subgroup scheme, which is multiplicative. Then*

$$\deg(\omega_{\mathcal{A}}) = \deg(\omega_{\mathcal{A}/G}).$$

*Proof.* By Lemma 4.11, we may assume that $G$ is diagonalisable. In particular, we may assume that there is a finite group scheme $G_0 \to \mathrm{Spec}(k)$ such that $G_{0,S} \simeq G$. Let $\mathcal{B} := \mathcal{A}/G$. Let $f : \mathcal{A} \to S$ and $g : \mathcal{B} \to S$

be the structural morphisms and let $\pi : \mathcal{A} \to \mathcal{B}$ be the quotient morphism. The triangle of cotangent complexes associated with the morphisms $\pi$, $g$ and $f$ gives an exact sequence

$$0 \to \mathcal{H}_1(\mathrm{CT}(\pi)) \to \pi^*(\Omega_g) \to \Omega_f \to \Omega_\pi \to 0, \tag{2}$$

where $\mathrm{CT}(\pi)$ is the cotangent complex of $\pi$ and $\mathcal{H}_1(\mathrm{CT}(\pi))$ is its first homology sheaf. Now $\pi$ makes $\mathcal{A}$ into a torsor over $\mathcal{B}$ and under $G_{\mathcal{B}}$. Hence there is a faithfully flat morphism $T \to \mathcal{B}$ (for instance, we may take $T = \mathcal{A}$), such that $\mathcal{A}_T \simeq (G_{\mathcal{B}}) \times_{\mathcal{B}} T$. In particular we have

$$\Omega_{\pi_T} \simeq \Omega_{G_0/k,T} \qquad \text{and} \qquad \mathcal{H}_1(\mathrm{CT}(\pi_T)) \simeq \mathcal{H}_1(\mathrm{CT}(G_0/k))_T$$

because the homology sheaves of the cotangent complex of $G_0$ over $k$ are flat (since they are $k$-vector spaces).

On the other hand, since $T \to \mathcal{B}$ is flat, we have

$$\Omega_{\pi_T} \simeq \Omega_{\pi,T} \qquad \text{and} \qquad \mathcal{H}_1(\mathrm{CT}(\pi_T)) \simeq \mathcal{H}_1(\mathrm{CT}(\pi))_T.$$

Finally, notice that $\Omega_{G_0/k,T}$ and $\mathcal{H}_1(\mathrm{CT}(G_0/k))_T$ are flat and thus by flat descent, the sheaves $\mathcal{H}_1(\mathrm{CT}(\pi))$ and $\Omega_\pi$ are flat (in other words: locally free). Hence the sequence

$$0 \to \epsilon^*_{\mathcal{A}/S}(\mathcal{H}_1(\mathrm{CT}(\pi))) \to \epsilon^*_{\mathcal{B}/S}(\Omega_g) \to \epsilon^*_{\mathcal{A}/S}(\Omega_f) \to \epsilon^*_{\mathcal{A}/S}(\Omega_\pi) \to 0 \tag{3}$$

is also exact. Furthermore, we then have

$$\epsilon^*_{\mathcal{A}/S}(\mathcal{H}_1(\mathrm{CT}(\pi))) \simeq \mathcal{H}_1(\mathrm{CT}(G_0/k))_S \qquad \text{and} \qquad \epsilon^*_{\mathcal{A}/S}(\Omega_\pi) \simeq \Omega_{G_0/k,S}$$

and thus the sheaves $\epsilon^*_{\mathcal{A}/S}(\mathcal{H}_1(\mathrm{CT}(\pi)))$ and $\epsilon^*_{\mathcal{A}/S}(\Omega_\pi)$ are trivial sheaves. In particular, we have that $\deg(\epsilon^*_{\mathcal{A}/S}(\mathcal{H}_1(\mathrm{CT}(\pi)))) = \deg(\epsilon^*_{\mathcal{A}/S}(\Omega_\pi)) = 0$ and by the additivity of $\deg(\cdot)$, we deduce from the existence of the sequence (3) that $\deg(\omega_{\mathcal{A}}) = \deg(\omega_{\mathcal{A}/G})$. □

**Remark 4.13.** The computation of the cotangent complex made in the proof of Lemma 4.11 is in essence also contained in [Ekedahl 1988, Proposition 1.1] (but the assumptions made there are not quite the right ones for us).

## 5. Proofs of the claims made in Section 2A

*We now use the terminology of the introduction.* So let $k$ be a finite field of characteristic $p > 0$ and let $S$ be a smooth, projective and geometrically connected curve over $k$. Let $K := \kappa(S)$ be its function field. Let $A$ be an abelian variety of dimension $g$ over $K$. Fix an algebraic closure $\overline{K}$ of $K$. Let $K^{\mathrm{perf}} \subseteq \overline{K}$ be the maximal purely inseparable extension of $K$ and let $K^{\mathrm{unr}} \subseteq K^{\mathrm{sep}}$ be the maximal separable extension of $K$, which is unramified above every place of $K$. Finally, we let $\mathcal{A}$ be a smooth commutative group scheme over $S$ such that $\mathcal{A}_K = A$.

*Proof of Theorem 2.1.* Recall the statement: there exists a (necessarily unique) multiplicative subgroup scheme $G_{\mathcal{A}} \hookrightarrow \ker F_{\mathcal{A}/S}$, with the following property: if $H$ is a multiplicative, finite and flat group

scheme of height one over $S$ and $f : H \to \ker F_{\mathcal{A}/S}$ is a morphism of group schemes, then $f$ factors through $G_{\mathcal{A}}$. If $A$ is ordinary and $\omega_{\mathcal{A}}$ is not ample then the order of $G_{\mathcal{A}}$ is $p^{\overline{\mathrm{rk}}_{\min}(\omega_{\mathcal{A}})}$. If $\phi : \mathcal{A} \to \mathcal{B}$ is a morphism of smooth commutative group schemes over $S$, then the restriction of $\phi$ to $G_{\mathcal{A}}$ factors through $G_{\mathcal{B}}$. Furthermore, we have $\deg(\omega_{\mathcal{A}}) = \deg(\omega_{\mathcal{A}/G_{\mathcal{A}}})$.

In spite of its lengthy statement, the proof Theorem 2.1 readily follows from Lemmata 4.8 and 4.12. More precisely, we simply have to define $G_{\mathcal{A}} := (\ker F_{\mathcal{A}/S})_{\mu}$ in the notation of Lemma 4.8. The equality $\deg(\omega_{\mathcal{A}}) = \deg(\omega_{\mathcal{A}/G_{\mathcal{A}}})$ now follows from Lemma 4.12. □

*Proof of Proposition 2.5.* Recall the assumptions of Proposition 2.5: $A$ is ordinary, $\mathcal{A}$ is semiabelian and $A(K^{\mathrm{perf}})$ is not finitely generated. We have to prove that $G_{\mathcal{A}}$ is of order $> 1$ and that $\mathcal{A}/G_{\mathcal{A}}$ is also semiabelian.

We know that $\bar{\mu}_{\min}(\omega_{\mathcal{A}/S}) \geq 0$ by Lemma 4.8 and since $A(K^{\mathrm{perf}})$ is not finitely generated, we know by Theorem B.1 that $\bar{\mu}_{\min}(\omega_{\mathcal{A}/S}) = 0$. Proposition 2.5 now follows from Theorem 2.1 and Lemma 4.10. □

*Proof of Proposition 2.6.* Recall the assumptions of Proposition 2.6: $A$ is ordinary, $\mathcal{A}$ is semiabelian over $S$ and $A(K^{\mathrm{perf}})$ is not finitely generated. We have to prove that there a finite flat morphism

$$\phi : \mathcal{A} \to \mathcal{B},$$

where $\mathcal{B}$ is a semiabelian over $S$ and a finite flat morphism

$$\lambda : \mathcal{B} \to \mathcal{B}$$

such that $\ker(\phi)$ are $\ker(\lambda)$ are multiplicative group schemes and such that the order of $\ker(\lambda)$ is $> 1$.

Consider now $\mathcal{A}_1 := \mathcal{A}/G_{\mathcal{A}}$. By Lemma 4.10, the group scheme $\mathcal{A}_1$ is also semiabelian and of course $A_1 := \mathcal{A}_{1,K}$ is also an ordinary abelian variety. We also have that $A_1(K^{\mathrm{perf}})$ is not finitely generated, since the natural map $A(K^{\mathrm{perf}}) \to A_1(K^{\mathrm{perf}})$ has finite kernel. Finally, the quotient morphism is $\mathcal{A} \to \mathcal{A}_1$ is finite, flat, with multiplicative kernel and $G_{\mathcal{A}}$ is nontrivial by Proposition 2.5.

Repeating the above procedure for $\mathcal{A}_1$ in place of $\mathcal{A}$ and continuing this way, we obtain an infinite sequence of semiabelian schemes over $S$

$$\mathcal{A} \to \mathcal{A}_1 \to \mathcal{A}_2 \to \cdots, \tag{4}$$

where all the connecting morphisms are finite, flat, of degree $> 1$ and with multiplicative kernel. Applying Lemma 4.12, we see that

$$\deg(\omega_{\mathcal{A}}) = \deg(\omega_{\mathcal{A}_1}) = \deg(\omega_{\mathcal{A}_2}) = \cdots$$

Let now $K'$ be a finite separable extension of $K$ such that $A(K)[n] \simeq (\mathbb{Z}/n\mathbb{Z})^{2\dim(A)}$ for some $n \geq 3$ such that $(p, n) = 1$. Let $S'$ be the normalisation of $S$ in $K'$. After base-change, we obtain an infinite sequence of semiabelian schemes over $S'$

$$\mathcal{A}_{S'} \to \mathcal{A}_{1,S'} \to \mathcal{A}_{2,S'} \to \cdots \tag{5}$$

and applying a theorem of Zarhin (see [Rössler 2013, Theorem 3.1] for a statement, explanations and further references), we conclude that in the sequence (5), there are only finitely many isomorphism

classes of semiabelian schemes over $S'$. On the other hand, applying a basic finiteness result in Galois cohomology proven by Borel and Serre (see [Zarkhin and Parshin 1989, paragraph 3, p. 69]), we can now conclude that in the sequence (4), there are also only finitely many isomorphism classes of semiabelian schemes over $S$.

Hence there are integers $j > i \geq 0$ and an isomorphism

$$I : \mathcal{A}_i \simeq \mathcal{A}_j$$

over $S$. Letting $\phi : \mathcal{A} \to \mathcal{A}_i$ be the constructed morphism and letting $\lambda$ be the constructed morphism $\mathcal{A}_i \to \mathcal{A}_j$ composed with $I^{-1}$, we can now conclude the proof of Proposition 2.6. □

## 6. Proofs of the claims made in Section 2B

**Theorem 2.7.** *Suppose that* $\mathrm{Tr}_{\overline{K}|\bar{k}}(A_{\overline{K}}) = 0$. *Suppose that the action of* $\mathrm{Gal}(K^{\mathrm{sep}}|K)$ *on* $\mathrm{Tor}_p(A(K^{\mathrm{unr}}))$ *factors through* $\mathrm{Gal}(K^{\mathrm{sep}}|K)^{\mathrm{ab}}$. *Then* $\mathrm{Tor}_p(A(K^{\mathrm{unr}}))$ *is finite.*

*Proof.* Let $L|K$ be the maximal subextension of $K^{\mathrm{unr}}|K$, which is Galois with abelian Galois group. Since $S$ is geometrically integral, $K \otimes_k \bar{k}$ is a field and $L$ contains a subfield isomorphic to $K \otimes_k \bar{k}$ (note that $\bar{k} = k^{\mathrm{sep}}$ and that $\mathrm{Gal}(\bar{k}|k) \simeq \widehat{\mathbb{Z}}$, which is an abelian group). Furthermore, geometric class field theory (see, e.g., [Szamuely 2010, Corollary 1.3]) tells us that $\mathrm{Gal}(L \mid K \otimes_k \bar{k})$ is a finite group. In particular, the field $L$ is finitely generated (as a field) over $\bar{k}$, since $K \otimes_k \bar{k}$ is finitely generated over $\bar{k}$. Now suppose to obtain a contradiction that $\mathrm{Tor}_p(A(K^{\mathrm{unr}}))$ were infinite. By assumption, we have

$$\mathrm{Tor}_p(A(K^{\mathrm{unr}})) \subseteq \mathrm{Tor}_p(A(L)).$$

Thus $\mathrm{Tor}_p(A(L))$ is infinite as well. By the Lang–Néron theorem, this implies that

$$\mathrm{Tr}_{L|\bar{k}}(A_L) \neq 0,$$

contradicting the first assumption. □

**Proposition 2.8.** *Suppose that* $\dim(A) \leqslant 2$ *and that* $\mathrm{Tr}_{\overline{K}|\bar{k}}(A_{\overline{K}}) = 0$. *Then* $\mathrm{Tor}_p(A(K^{\mathrm{unr}}))$ *is finite.*

*Proof.* Notice that if $\mathrm{Tor}_p(A(K^{\mathrm{unr}}))$ is infinite then we have

$$\bigcap_{\ell \geq 0} p^\ell \cdot \mathrm{Tor}_p(A(K^{\mathrm{unr}})) \neq 0$$

This follows from the fact that for each $n \geq 0$, the set

$$\{x \in \mathrm{Tor}_p(A(K^{\mathrm{unr}})) \mid p^n \cdot x = 0\}$$

is finite (the details are left to the reader). Let $G \subseteq \bigcap_{\ell \geq 0} p^\ell \cdot (\mathrm{Tor}_p(A(K^{\mathrm{unr}})))$ be the subgroup of elements annihilated by the multiplication by $p$ map.

If $G = 0$ then there the conclusion holds, because then $\bigcap_{\ell \geq 0} p^\ell \cdot (\mathrm{Tor}_p(A(K^{\mathrm{unr}}))) = 0$ and thus $\mathrm{Tor}_p(A(K^{\mathrm{unr}}))$ is finite by the above remark.

Suppose now that $\#G = p$. Then $\bigcap_{\ell \geq 0} p^\ell \cdot \mathrm{Tor}_p(A(K^{\mathrm{unr}}))$ is infinite and it is a union of cyclic groups of $p$-power order (use the classification theorem for finite abelian groups). Thus the action of $\mathrm{Gal}(K^{\mathrm{sep}}|K)$ on $\bigcap_{\ell \geq 0} p^\ell \cdot (\mathrm{Tor}_p(A(K^{\mathrm{unr}})))$ factors through $\mathrm{Gal}(K^{\mathrm{sep}}|K)^{\mathrm{ab}}$. But this contradicts Theorem 2.7 and thus we must have $\#G > p$. If $\#G > p$ then by the assumption that $\dim(A) \leq 2$, we see that we must have $\#G = p^2$ and thus the inclusions

$$\mathrm{Tor}_p(A(K^{\mathrm{unr}})) \subseteq \mathrm{Tor}_p(A(K^{\mathrm{sep}})) \subseteq \mathrm{Tor}_p(A(\overline{K}))$$

are both equalities. In particular, $A$ is an ordinary abelian surface. Let now $s \in S$ be a closed point such that $\mathcal{A}_s$ is an ordinary abelian variety over $s$. Let $W := \mathrm{Spec}(\widehat{\mathcal{O}^{\mathrm{sh}}_{S,s}})$ be the spectrum of the completion of the strict henselisation of the local ring at $s$ and write $\widehat{K^{\mathrm{sh}}_s}$ for the fraction field of $\widehat{\mathcal{O}^{\mathrm{sh}}_{S,s}}$. The abelian scheme $\mathcal{A}_W \to W$ gives rise to an element $e$ of

$$\mathrm{Hom}_{\mathbb{Z}_p}(T_p(\mathcal{A}_{\bar{s}}(\bar{s})) \otimes T_p(\mathcal{A}^\vee_{\bar{s}}(\bar{s})), \widehat{\mathcal{O}^{\mathrm{sh}}_{S,s}}^{\,*}).$$

Here $T_p(\mathcal{A}_{\bar{s}}(\bar{s}))$ and $T_p(\mathcal{A}^\vee_{\bar{s}}(\bar{s}))$ are the $p$-adic Tate modules of $\mathcal{A}_{\bar{s}}$ and $\mathcal{A}^\vee_{\bar{s}}$ respectively and $\widehat{\mathcal{O}^{\mathrm{sh}}_{S,s}}^{\,*}$ is the group of multiplicative units of $\widehat{\mathcal{O}^{\mathrm{sh}}_{S,s}}$. The element $e$ is called the Serre–Tate pairing associated with $\mathcal{A}_W$. See [Katz 1981] for the construction of this pairing. We have $e = 0$ if and only if $\mathcal{A}_W \simeq \mathcal{A}_{\bar{s}} \times_{\bar{s}} W$. Furthermore, the fact that

$$\mathrm{Tor}_p(\mathcal{A}(W)) = \mathrm{Tor}_p(A(\widehat{K^{\mathrm{sh}}_s})) = \mathrm{Tor}_p(A(K^{\mathrm{unr}})) = \mathrm{Tor}_p(A(\overline{\widehat{K^{\mathrm{sh}}_s}}))$$

in our situation shows that $e = 0$. This follows directly from the definition of the Serre–Tate pairing in the ordinary case (see the definition of the morphism "$p^n$" in [Katz 1981, p. 151]). Thus we have $\mathcal{A}_W \simeq \mathcal{A}_{\bar{s}} \times_{\bar{s}} W$ and in particular $\mathrm{Tr}_{\overline{K}|\bar{k}}(A_{\overline{K}}) \neq 0$ by Proposition 9.1 (c) below. This contradicts one of our assumptions. We conclude that $G = 0$, so the conclusion must hold. □

**Theorem 2.9.** *Suppose that* $\mathrm{Tor}_p(A(K^{\mathrm{sep}}))$ *is infinite. Then there is an étale* $K$*-isogeny*

$$\phi : A \to B,$$

*where* $B$ *is an abelian variety over* $K$ *and there is an étale* $K$*-isogeny*

$$\lambda : B \to B$$

*such that the order of* $\ker(\lambda)$ *is* $> 1$ *and such that the orders of* $\ker(\lambda)$ *and* $\ker(\phi)$ *are powers of* $p$*.*

*Proof.* Note that in [Rössler 2013, Theorem 1.4], this statement is proven under the supplementary assumption that there exist $n \in \mathbb{Z}$, such that $(n, p) = 1$ and $n > 3$ and such that $A[n](\overline{K}) \simeq (\mathbb{Z}/n\mathbb{Z})^{2\dim(A)}$. Using [Zarkhin and Parshin 1989, paragraph 3, "Finiteness Theorem for Forms", p. 69] in the proof, it can be seen that this assumption is not necessary. A completely parallel argument is described in the proof of Proposition 2.6. We leave the details to the reader. □

**Theorem 2.10.** *Suppose that there exists an étale* $K$*-isogeny* $\phi : A \to A$*, such that* $\deg(\phi)$ *is strictly larger than* 1 *and that* $\deg(\phi) = p^r$ *for some* $r > 0$*. Suppose also that* $A$ *is a geometrically simple abelian variety and that* $\mathcal{A}$ *is a semiabelian scheme.*

*Then* $\mathcal{A}$ *is an abelian scheme and* $\phi$ *extends to an étale* $S$*-morphism* $\mathcal{A} \to \mathcal{A}$ *of group schemes.*

*Proof.* Notice first that by a result of Raynaud [1970, IX, Corollary 1.4, p. 130], the morphism $\phi$ extends uniquely to an $S$-morphism $\bar{\phi} : \mathcal{A} \to \mathcal{A}$ of group schemes. Since $\bar{\phi}$ is étale over $K$, we have an exact sequence of coherent sheaves

$$0 \to \bar{\phi}^*(\Omega_{\mathcal{A}/S}) \to \Omega_{\mathcal{A}/S}$$

on $\mathcal{A}$. Let $\sigma \in H^0(\mathcal{A}, \det(\bar{\phi}^*(\Omega_{\mathcal{A}/S}))^\vee \otimes \det(\Omega_{\mathcal{A}/S}))$ be the corresponding section. Since

$$\sigma_K \in H^0(A, \det(\phi^*(\Omega_{A/K}))^\vee \otimes \det(\Omega_{A/K}))$$

has an empty zero-scheme, the zero scheme $Z(\sigma)$ is supported on a finite number of closed fibres of $\mathcal{A}$. Hence there exists a finite number $P_1, \ldots, P_n$ of closed points of $S$, such that $Z(\sigma) = \coprod_{i=1}^n n_i \mathcal{A}_{P_i}$ (as Weil divisors) for some $n_i \geqslant 0$. On the other hand, the Weil divisor $Z(\sigma)$ is rationally equivalent to 0, since $\det(\phi^*(\Omega_{\mathcal{A}/S}))^\vee \otimes \det(\Omega_{\mathcal{A}/S}) \simeq \det(\Omega_{\mathcal{A}/S})^\vee \otimes \det(\Omega_{\mathcal{A}/S}) \simeq \mathcal{O}_{\mathcal{A}}$. Now notice that the morphism $p^* : \mathrm{Pic}(S) \to \mathrm{Pic}(\mathcal{A})$ of Picard groups is injective, because it is split by the map $\epsilon^*_{\mathcal{A}/S} : \mathrm{Pic}(\mathcal{A}) \to \mathrm{Pic}(S)$. Hence the Weil divisor $\coprod_{i=1}^n n_i P_i$ is rationally equivalent to 0 on $S$, which implies that $n_i = 0$ for all $i = 1, \ldots n$. In other words, we have $Z(\sigma) = \varnothing$ and thus the morphism $\bar{\phi}^*(\Omega_{\mathcal{A}/S}) \to \Omega_{\mathcal{A}/S}$ is an isomorphism. By [Hartshorne 1977, III, Proposition 10.4], this implies that $\bar{\phi}$ is étale.

Let now $s \in S$ be a closed point such that $\mathcal{A}_s$ has a presentation

$$0 \to G \xrightarrow{\iota} \mathcal{A}_s \to A_0^0 \to 0,$$

where $G$ is a torus over $s$ of dimension $d > 0$ and $A_0^0$ is an abelian variety over $s$. The morphism $\bar{\phi}_s|_G : G \to \mathcal{A}_s$ factors through $G$, since there is no nonconstant $s$-morphism $G \to A_0^0$. Call $\gamma : G \to G$ the resulting morphism. The morphism $\gamma$ is étale. Indeed, we have a commutative diagram

$$\begin{array}{ccccc}
\gamma^*(\iota^*(\Omega_{\mathcal{A}_s/s})) & \longrightarrow & \gamma^*(\Omega_{G/s}) & \longrightarrow & \Omega_{G/s} \\
\downarrow \sim & & & & \downarrow = \\
\iota^*(\bar{\phi}_s^*(\Omega_{\mathcal{A}_s/s})) & \longrightarrow & \iota^*(\Omega_{\mathcal{A}_s/s}) & \longrightarrow & \Omega_{G/s}
\end{array}$$

and in the lower row of this diagram all the arrows are surjective. Thus the arrow

$$\gamma^*(\Omega_{G/s}) \to \Omega_{G/s}$$

must also be surjective and hence an isomorphism. Since $G$ is smooth over $\kappa(s)$, we conclude that $\gamma$ is smooth by [Hartshorne 1977, III, Proposition 10.4]. In particular $\gamma$ is faithfully flat, because it is a morphism of group schemes and $G$ is connected (see, e.g., [SGA $3_1$ 2011, Exposé IV-B, Corollary 1.3.2]). Now recall that there is a $K$-morphism $\psi : A \to A$ such that $\psi \circ \phi = [p^{\deg(\phi)}]_A$ (because finite commutative group schemes over $K$ are annihilated by their order; see [Tate and Oort 1970, Theorem (Deligne), p. 4]). The morphism $\psi$ extends uniquely to $\bar{\psi} : \mathcal{A} \to \mathcal{A}$ and thus by unicity, we have $\bar{\psi} \circ \bar{\phi} = [p^{\deg(\phi)}]_{\mathcal{A}}$. In particular, $\ker(\gamma)$ is a closed subscheme of $\ker([p^{\deg(\phi)}]_G)$. Since $\ker([p^{\deg(\phi)}]_G)$ is an infinitesimal group scheme and $\gamma$ is étale, we see that $\ker(\gamma) = 0$ (since $\ker(\gamma)$ is étale over $s$). Thus $\gamma$ is an isomorphism.

Now choose a $\bar{s}$-isomorphism $G_{\bar{s}} \simeq \mathbb{G}_m^d$ (here $\bar{s}$ if the spectrum of the algebraic closure of $\kappa(s)$). The morphism $\gamma_{\bar{s}}$ is described by a matrix $M \in \mathrm{GL}_d(\mathbb{Z})$ (because the group scheme dual to $G_{\bar{s}}$ is the diagonalisable group scheme over $\bar{s}$ associated with $\mathbb{Z}^d$). Hence there exists a monic polynomial $P(x) \in \mathbb{Z}[x]$, such that $P(0) = \pm 1$ and such that $P(\gamma_{\bar{s}}) = 0$.

Finally, choose a prime $l \neq p$. Let $\widehat{\mathcal{O}}_s^{\mathrm{sh}}$ be the completion of the strict henselisation of the local ring of $S$ at $s$. Let $\widehat{K}_s^{\mathrm{sh}}$ be the fraction field of $\widehat{\mathcal{O}}_s^{\mathrm{sh}}$ and let $j \in \mathbb{N}$. The closed subgroup scheme $G_{\bar{s}}[l^j]$ of $G_{\bar{s}}$ extends uniquely to a finite and étale subgroup scheme $\widetilde{G}_{l^j}$ of $\mathcal{A}_{\widehat{\mathcal{O}}_s^{\mathrm{sh}}}$ over $\widehat{\mathcal{O}}_s^{\mathrm{sh}}$. See [SGA $3_{\mathrm{II}}$ 1970, Theorem 3.6 and Theorem 3.6 bis]. Furthermore the natural map $\widetilde{G}_{l^j}(\widehat{\mathcal{O}}_s^{\mathrm{sh}}) \to G_{\bar{s}}[l^j](\bar{s})$ is a bijection, since $\widehat{\mathcal{O}}_s^{\mathrm{sh}}$ is strictly henselian and $\widetilde{G}_{l^j}$ is étale (see [Milne 1980, Proposition I.4.4]). Hence $P(\phi)(\widetilde{G}_{l^j}(\widehat{\mathcal{O}}_s^{\mathrm{sh}})) = 0$. On the other hand, the image of the group $\bigcup_{j\in\mathbb{N}} \widetilde{G}_{l^j}(\widehat{\mathcal{O}}_s^{\mathrm{sh}})$ in $A_{\widehat{K}_s^{\mathrm{sh}}}$ is dense, because $A$ is geometrically simple and the group $\bigcup_{j\in\mathbb{N}} \widehat{G}_{l^j,\bar{s}}(\mathcal{O}_s^{\mathrm{sh}})$ is infinite. Hence $P(\phi) = 0$ and since $P(0) = \pm 1$, we see that $\phi$ is an automorphism, which is a contradiction. □

## 7. Proof of Theorem 1.1

**Theorem 1.1.** (a) *Suppose that $A$ is geometrically simple. If $A(K^{\mathrm{perf}})$ is finitely generated and of rank $> 0$ then $\mathrm{Tor}_p(A(K^{\mathrm{sep}}))$ is a finite group.*

(b) *Suppose that $A$ is an ordinary (not necessarily simple) abelian variety. If $\mathrm{Tor}_p(A(K^{\mathrm{sep}}))$ is a finite group then $A(K^{\mathrm{perf}})$ is finitely generated.*

We shall need the following:

**Lemma 7.1.** *Let $B$ be an abelian variety over $K$ and let $\gamma : B \to B$ be a $K$-isogeny such that $\deg(\phi) > 1$. Suppose that $B$ is geometrically simple. Let $H \subseteq A(\overline{K})$ be a finitely generated subgroup. Then the set*

$$\bigcap_{r\geq 0} \gamma^{\circ r}(H)$$

*is a finite group.*

*Proof of Lemma 7.1.* Let $G := \bigcap_{r\geq 0} \gamma^{\circ r}(H)$. Let $F := G/\mathrm{Tor}(G)$ be the quotient of $G$ by its torsion subgroup. We may suppose without restriction of generality that $\mathrm{rk}(G) > 0$ for otherwise the lemma is proven. Since $\gamma$ is a group homomorphism, we have $\gamma(\mathrm{Tor}(G)) \subseteq \mathrm{Tor}(G)$ and thus $\gamma$ gives rise to a group homomorphism $F \to F$ that we also denote by $\gamma$. By construction, we have $\gamma(F) = F$ and thus $\gamma : F \to F$ is a bijection, since $F$ is a finitely generated free $\mathbb{Z}$-module. Let

$$P(t) := t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0 \in \mathbb{Z}[t]$$

be the characteristic polynomial of $\gamma : F \to F$. We have $P(\gamma) = 0$ by the Cayley–Hamilton theorem and since $\gamma$ is an automorphism, we have

$$P(0) = a_0 = \pm 1 = \det(\gamma).$$

Hence

$$(-a_0)^{-1}\cdot(\gamma^{\circ,n-1}+a_{n-1}\cdot\gamma^{\circ,n-2}+\ldots a_1\cdot\mathrm{Id}_F)$$

is the inverse of $\gamma : F \to F$. Now let $\widetilde{\gamma}$ be the $K$-group scheme homomorphism

$$\widetilde{\gamma} := (-a_0)^{-1}\cdot(\gamma^{\circ,n-1}+a_{n-1}\cdot\gamma^{\circ,n-2}+\ldots a_1\cdot\mathrm{Id}_B)$$

from $B$ to $B$. Suppose first that the morphism of $K$-group schemes $\widetilde{\gamma}\circ\gamma-\mathrm{Id}_B$ is not the zero morphism. Then it is surjective, because $B$ is simple. Furthermore the group $G$ is dense in $B_{\overline{K}}$, since $B$ is geometrically simple. Thus the group $(\widetilde{\gamma}\circ\gamma-\mathrm{Id}_B)(G)$ is dense in $B_{\overline{K}}$. On the other hand, by construction $(\widetilde{\gamma}\circ\gamma-\mathrm{Id}_B)(G)\subseteq\mathrm{Tor}(G)$. Since $\mathrm{Tor}(G)$ is a finite group, it is not dense in $B_{\overline{K}}$ and thus we deduce that $\widetilde{\gamma}\circ\gamma-\mathrm{Id}_B$ must be the zero morphism. Hence $\gamma$ is invertible (with inverse $\widetilde{\gamma}$), which contradicts the assumption that $\deg(\gamma)>1$. We conclude that we cannot have $\mathrm{rk}(G)>0$ and thus $G=\mathrm{Tor}(G)$ is a finite group. □

*Proof of Theorem 1.1.* For statement (a), suppose first that $\mathrm{Tor}_p(A(K^{\mathrm{sep}}))$ is not a finite group. Then by Theorem 2.9, there exists an abelian variety $B$ over $K$, which is $K$-isogenous to $A$ and which carries an étale $K$-endomorphism $B\to B$, whose degree is $>1$ and is a power of $p$. The dual of $B$ hence carries an isogeny $\phi$, which is purely inseparable (because the dual of a finite étale group scheme over a field is an infinitesimal group scheme) and thus we have

$$B^{\vee}(K^{\mathrm{perf}})=\bigcap_{r\geq 0}\phi^{\circ r}(B^{\vee}(K^{\mathrm{perf}}))$$

By Lemma 7.1, $B^{\vee}(K^{\mathrm{perf}})$ is thus either finite or not finitely generated and the same holds for $A$, since $A$ is isogenous to $B^{\vee}$. This proves (a).

We now turn to the proof of statement (b). Note that by Grothendieck's semiabelian reduction theorem, there is a finite and separable extension $K_1|K$ such that $A_{K_1}$ extends to a semiabelian scheme over the normalisation $S_1$ of $S$ in $K_1$. The scheme $S_1$ might not be geometrically connected over $k$ but there a finite extension $k_1$ of $k$, such that the connected components of $S_{1,k_1}$ are geometrically connected. We choose one of these connected components, say $S_2$. The extension of function fields corresponding to the morphism $S_2\to S$ is separable by construction so we may (and do) assume that $\mathcal{A}$ is semiabelian to begin with. Suppose that $A(K^{\mathrm{perf}})$ is not finitely generated and that $A$ is ordinary. Then by Proposition 2.6, there is an abelian variety $B$ over $K$, which is $K$-isogenous to $A$ and which carries a $K$-isogeny $B\to B$, whose kernel is a multiplicative group scheme of order $>1$. The dual $\phi$ of this isogeny is an étale isogeny of $B^{\vee}$, which has degree $p^r$ for some $r>0$. Thus $\mathrm{Tor}_p(B^{\vee}(K^{\mathrm{sep}}))$ is an infinite group and the same holds for $A$, since $A$ is isogenous to $B^{\vee}$. This proves (b). □

## 8. Proof of Theorem 1.2

**Theorem 1.2.** *Suppose that $\mathcal{A}$ is a semiabelian scheme and that $A$ is a geometrically simple abelian variety over $K$. If $\mathrm{Tor}_p(A(K^{\mathrm{sep}}))$ is infinite, then*

(a) *$\mathcal{A}$ is an abelian scheme*;

(b) *there is $r_A\geq 0$ such that $p^{r_A}\cdot\mathrm{Tor}_p(A(K^{\mathrm{sep}}))\subseteq\mathrm{Tor}_p(A(K^{\mathrm{unr}}))$.*

*Furthermore, there is*

(c) *an abelian scheme $\mathcal{B}$ over $S$;*

(d) *a generically étale $S$-isogeny $\mathcal{A} \to \mathcal{B}$, whose degree is a power of $p$;*

(e) *an étale $S$-isogeny $\mathcal{B} \to \mathcal{B}$ whose degree is $> 1$ and is a power of $p$.*

*Finally*

(f) *if $A$ is ordinary then the Kodaira–Spencer rank of $A$ is not maximal;*

(g) *if* $\dim(A) \leqslant 2$ *then* $\mathrm{Tr}_{\bar{K}|\bar{k}}(A_{\bar{K}}) \neq 0$.

(h) *for all closed points $s \in S$, the $p$-rank of $\mathcal{A}_s$ is $> 0$.*

*Proof.* Proof of (a): Note that by Theorem 2.9, the abelian variety $A$ is isogenous to an abelian variety $B$ over $K$, which is endowed with an étale endomorphism of degree a positive power of $p$. Since $A$ extends to a semiabelian scheme over $S$ so does $B$. This is a consequence of a theorem of Grothendieck (see [Abbes 2000, 5.] for a nice presentation). Thus, by Theorem 2.10 we see that $B$ extends to an abelian scheme $\mathcal{B}$ over $S$. Using the criterion of Néron–Ogg–Shafarevich (see [Serre and Tate 1968]), we see that $A$ also extends to an abelian scheme over $S$. By the uniqueness of semiabelian models (see [Raynaud 1970, IX, Corollary 1.4, p. 130]), this extension must be $\mathcal{A}$ and thus $\mathcal{A}$ is an abelian scheme.

Proof of (b): Let $H := \mathrm{Gal}(K^{\mathrm{sep}}|K^{\mathrm{unr}})$. For $i \geq 0$, let $G_i := A(K^{\mathrm{sep}})[p^i]$. The group $G_i$ is the group of $K$-rational points of an étale finite group scheme $\underline{G}_i$ over $K$, which is naturally a closed subgroup scheme of $A$. Let $A_i := A/\underline{G}_i$ and for $i \leq j$ let $\phi_{i,j} : A_i \to A_j$ be the natural morphism. Let $\mathcal{A}_i$ be the connected component of the zero section of the Néron model of $A_i$ over $S$. By (a) and the criterion of Néron–Ogg–Shafarevich (see [Serre and Tate 1968]), this is an abelian scheme. Furthermore, by [Raynaud 1970, IX, Corollary 1.4, p. 130] the morphisms $\phi_{i,j}$ extend to morphisms $\bar{\phi}_{i,j} : \mathcal{A}_i \to \mathcal{A}_j$ and we have the classical exact sequence

$$\bar{\phi}_{i,j}^*(\Omega_{\mathcal{A}_j/S}) \to \Omega_{\mathcal{A}_i/S} \to \Omega_{\bar{\phi}_{i,j}} \to 0.$$

Now the morphism $\bar{\phi}_{i,j}^*(\Omega_{\mathcal{A}_j/S}) \to \Omega_{\mathcal{A}_i/S}$ is injective over the generic point of $\mathcal{A}_i$, because $\phi_{i,j} = \bar{\phi}_{i,j,K}$ is smooth by construction. On the other hand both $\bar{\phi}_{i,j}^*(\Omega_{\mathcal{A}_j/S})$ and $\Omega_{\mathcal{A}_i/S}$ are locally free and thus it follows that $\bar{\phi}_{i,j}^*(\Omega_{\mathcal{A}_j/S}) \to \Omega_{\mathcal{A}_i/S}$ is also injective. Hence we have an exact sequence

$$0 \to \bar{\phi}_{i,j}^*(\Omega_{\mathcal{A}_j/S}) \to \Omega_{\mathcal{A}_i/S} \to \Omega_{\bar{\phi}_{i,j}} \to 0. \tag{6}$$

Let $\pi_i : \mathcal{A}_i \to S$ be the structural morphism. We have a functorial isomorphism

$$\Omega_{\mathcal{A}_i} \simeq \pi_i^*(\pi_{i,*}(\Omega_{\mathcal{A}_i/S}))$$

and thus there is a coherent sheaf $T_{i,j}$ on $S$, which is a torsion sheaf, such that $\pi_i^*(T_{i,j}) \simeq \Omega_{\bar{\phi}_{i,j}}$ and the sequence (6) is the pull-back by $\pi_i^*$ of a sequence

$$0 \to \pi_{j,*}(\Omega_{\mathcal{A}_j/S}) \to \pi_{i,*}(\Omega_{\mathcal{A}_i/S}) \to T_{i,j} \to 0$$

and in particular

$$\deg_S(\pi_{j,*}(\Omega_{\mathcal{A}_j/S})) + \deg_S(T_{i,j}) = \deg_S(\pi_{i,*}(\Omega_{\mathcal{A}_i/S})).$$

Now recall that $\deg_S(\pi_{i,*}(\Omega_{\mathcal{A}_i/S})) \geqslant 0$ for all $i \geq 0$ (see [Faltings and Chai 1990, V, Proposition 2.2, p. 164]). Thus, for $i = 0, 1, \ldots$, the sequence $\deg_S(\pi_{i,*}(\Omega_{\mathcal{A}_i/S}))$ is a nonincreasing sequence of natural numbers. Hence for large enough $i$, say $i_0$, it reaches its minimum. We conclude that $T_{i_0,j} = 0$ for $j > i_0$, so that the morphism $\bar{\phi}_{i_0,j}$ is étale and finite. Now $\phi_{0,i_0}(G_j(K^{\mathrm{sep}}))$ lies by construction in the kernel of $\phi_{i_0,j}$. Thus

$$\phi_{0,i_0}(G_j(K^{\mathrm{sep}})) \subseteq A_{i_0}(K^{\mathrm{unr}})$$

when $j > i_0$. In other words, for any $x \in G_j(K^{\mathrm{sep}})$ and any $\gamma \in H$, we have

$$\gamma(x) - x \in G_{i_0}(K^{\mathrm{sep}}).$$

In particular, we have

$$\gamma(p^{i_0} \cdot x) = p^{i_0} \cdot \gamma(x) = p^{i_0} \cdot x$$

In particular, since $j > i_0$ was arbitrary, we see that

$$\gamma(p^{i_0} \cdot x) = p^{i_0} \cdot x$$

for all $x \in \mathrm{Tor}_p(A(K^{\mathrm{sep}}))$ and all $\gamma \in H$. Setting $r_A = i_0$ concludes the proof of (b).

Proof of the existence statements (c), (d), (e): this is a consequence of (a) and Theorems 2.9 and 2.10.

Proof of (f): this is contained in a theorem of J.-F. Voloch; see [Voloch 1995, Proposition on p. 1093].

Proof of (g): this is a consequence of (b) and Proposition 2.8.

Proof of (h): This follows from (a) and (e). □

## 9. Proof of Theorem 1.4

**9A.** ***The trace of an abelian variety over a function field: basic facts.*** Let $E$ be an abelian over a field $F$. Let $F_0 \subseteq F$ be a subfield.

The $F|F_0$ trace $(\mathrm{Tr}_{F|F_0}(E), \lambda)$ (if it exists) of $E$ over $F_0$ is an abelian variety $\mathrm{Tr}_{F|F_0}(E)$ over $F_0$ together with a homomorphism $\lambda : \mathrm{Tr}_{F|F_0}(E)_F \to E$ of abelian varieties over $F$. They have the following universal property. For any abelian $E_0$ over $F_0$ and a homomorphism $\phi : E_{0,F} \to E$ of abelian varieties, there is a unique morphism $\widetilde{\phi} : E_{0,F} \to \mathrm{Tr}_{F|F_0}(E)_F$ such that $\phi = \lambda \circ \widetilde{\phi}$. This means that $\mathrm{Tr}_{F|F_0}(E)$ and $\lambda$ are uniquely determined if they exist.

Here are some known facts about $\mathrm{Tr}_{F|F_0}(E)$. Before stating them, we record the fact for any finite morphism of abelian varieties $f : E' \to E$ over $F$, the natural morphism $E'/\ker(f) \to E$ is a closed immersion. Here $E'/\ker(f)$ is the quotient described in Proposition 4.1. To see this, consider that the morphism $E'/\ker(f) \to E$ is by definition a monomorphism of fppf sheaves over $F_0$ and hence a monomorphism of schemes. On the other hand, it is proper and of finite type and thus a closed immersion (see [EGA $\mathrm{IV}_4$ 1967, p. 182] for this). We shall call $\mathrm{Im}(f)$ the abelian variety $E'/\ker(f)$ viewed as an abelian subvariety of $E$.

The field extension $F|F_0$ is called primary (resp. regular) if the algebraic closure of $F_0$ in $F$ is purely inseparable over $F_0$ (if $F_0$ is algebraically closed in $F$ and $F$ is separable over $F_0$). Note that if $F$ is the function field of a smooth and geometrically integral variety over $F_0$ then $F|F_0$ is regular.

**Proposition 9.1** (see [Conrad 2006, Theorems 6.4 and 6.12]). (a) *If $F|F_0$ is primary then the $F|F_0$ trace $(\mathrm{Tr}_{F|F_0}(E), \lambda)$ of $E$ over $F_0$ exists and the kernel of $\lambda$ is finite over $F$.*

(b) *If $F|F_0$ is regular then the kernel of the morphism $\lambda$ is connected and so is its Cartier dual.*

(c) *If $F_1|F$ and $F|F_0$ are primary extensions then $(\mathrm{Tr}_{F|F_0}(E)_{F_1}, \lambda_{F_1})$ is an $F_1|F_0$-trace of $E_{F_1}$.*

(d) *We have $\mathrm{Tr}_{F|F_0}(A/\mathrm{Im}(\lambda)) = 0$.*

We also recall the *Lang–Néron theorem* (see [Conrad 2006, Theorem 7.1; Lang 1983, Chapter 6, Theorem 2]): if $F|F_0$ is a finitely generated regular extension then the quotient group $E(F)/\mathrm{Tr}_{F|F_0}(E)(F_0)$ is finitely generated. Here $\mathrm{Tr}_{F|F_0}(E)(F_0)$ is viewed as a subgroup of $E(F)$ via $\lambda$ and the natural base change map from $F_0$ to $F$.

**9B.** ***The proof.*** We now use the notations of Conjecture 1.3.

Let $\lambda : \mathrm{Tr}_{L|l_0}(C) \to C$ be the canonical morphism. We write $C/\mathrm{Im}(\lambda)$ for the quotient of $C$ by $\mathrm{Im}(\lambda)$ in the sense of Proposition 4.1.

We begin with:

**Proposition 9.2.** *If* $\mathrm{IVD}(C/\mathrm{Im}(\lambda), L) \subseteq \mathrm{Tor}^p((C/\mathrm{Im}(\lambda))(L))$ *then* $\mathrm{IVD}(C, L) \subseteq \mathrm{Tor}^p(C(L))$.

For the proof of Proposition 9.2, we shall need the following:

**Lemma 9.3.** *Let $N$ be a finite flat infinitesimal group scheme over a field $J$ of characteristic $p$. There is a finite field extension $J'|J$ such that for any $n \geq 0$ and any element $\alpha \in H^1(J, N^{(p^n)})$, the image $\alpha_{J'}$ of $\alpha$ in $H^1(J', N^{(p^n)}{}_{J'})$ vanishes.*

Here $H^1(J, N^{(p^n)})$ is the first cohomology group of $N^{(p^n)}$ viewed as a sheaf in the fppf topology. More concretely, it is the group of isomorphism classes of torsors of $N^{(p^n)}$ over $J$. In the following proof, we shall write $J^{p^{-m}} \subseteq \bar{J}$ for the subfield of $\bar{J}$ consisting of elements of the form $x^{p^{-m}}$, where $x \in J$.

*Proof of Lemma 9.3.* First suppose that $N$ has a filtration by finite closed subgroup schemes, whose quotients are isomorphic to either $\alpha_{p,J}$ or $\mu_{p,J}$. Let $m \geq 0$ be the number of nonvanishing quotients. We shall prove by induction on $m$ that the image of $\alpha$ in $H^1(J^{p^{-m}}, N^{(p^n)})$ vanishes for all $n \geq 0$ (under the supplementary assumption on $N$), for any field $J$ of characteristic $p$. If $m = 0$ the statement holds tautologically, so we shall suppose that it holds for $1, \ldots, m-1$. Let

$$0 \to F_1 \to N_{J_1} \to F_2 \to 0$$

be a presentation of $N$ where $F_2$ is isomorphic to either $\alpha_{p,J}$ or $\mu_{p,J}$ and $F_1$ has a filtration as above, whose number of nonvanishing quotients is $\leq m-1$. This induces exact sequences

$$0 \to H^1(J^{p^{-1}}, (F_{1,J^{p^{-1}}})^{(p^n)}) \to H^1(J^{p^{-1}}, (N_{J^{p^{-1}}})^{(p^n)}) \to H^1(J^{p^{-1}}, (F_{2,J^{p^{-1}}})^{(p^n)})$$
$$0 \to H^1(J^{p^{-m}}, (F_{1,J^{p^{-m}}})^{(p^n)}) \to H^1(J^{p^{-m}}, (N_{J^{p^{-m}}})^{(p^n)}) \to H^1(J^{p^{-m}}, (F_{2,J^{p^{-m}}})^{(p^n)})$$

(observe that $H^0(J^{p^{-m}}, (F_{2,J^{p^{-m}}})^{(p^n)}) = 0$ since $F_2$ is infinitesimal). Since $F_2{}^{(p^n)}$ is of height one, the image of $\alpha$ in $H^1(J^{p^{-1}}, (F_{2,J^{p^{-1}}})^{(p^n)})$ vanishes by [Milne 2006, Lemma III.3.5.7]. The element $\alpha$ is thus

the image of an element $\beta \in H^1(J^{p^{-1}}, (F_{1,J^{p^{-1}}})^{(p^n)})$. By the inductive hypothesis, the image of $\beta$ in $H^1(J^{p^{-m}}, (F_{1,J^{p^{-m}}})^{(p^n)})$ vanishes and thus the image of $\alpha$ in $H^1(J^{p^{-m}}, (N_{J^{p^{-m}}})^{(p^n)})$ vanishes, proving the claim.

Now according to [Grothendieck 1974, §2.4, p. 28] there is a finite extension $J_1$ of $J$ such that $N_{J_1}$ has a filtration by finite closed subgroup schemes, whose quotients are isomorphic to either $\alpha_{p,J_1}$ or $\mu_{p,J_1}$. This extension will by construction also work for all the group schemes $N^{(p^n)}$ and the number of nonvanishing quotients of all the group schemes $N^{(p^n)}{}_{J_1}$ is constant, say it is $m$. Hence the extension $J' := J_1^{p^{-m}}$ has the required property. □

*Proof of Proposition 9.2.* . Now suppose that $\mathrm{IVD}(C/\mathrm{Im}(\lambda), L) \subseteq \mathrm{Tor}^p((C/\mathrm{Im}(\lambda))(L))$. We want to show that $\mathrm{IVD}(C, L) \subseteq \mathrm{Tor}^p(C(L))$.

Write

$$\lambda^{(p^n)} : \mathrm{Tr}_{L|l_0}(C)^{(p^n)} \to C^{(p^n)}$$

for the base change of $\lambda$ by $F_L^{\circ n}$. We have an exact sequence

$$0 \to \mathrm{Im}(\lambda)(L) \to C(L) \to (C/\mathrm{Im}(\lambda))(L)$$

and we have $(C/\mathrm{Im}(\lambda))^{(p^n)} \simeq C^{(p^n)}/\mathrm{Im}(\lambda^{(p^n)})$. Let now

$$x_0 \in C(L), \quad x_1 \in C^{(p)}(L), \quad x_2 \in C^{(p^2)}(L), \quad \ldots,$$

be a sequence of points such $V_{C^{(p)}/L}(x_1) = x_0$, $V_{C^{(p^2)}/L}(x_2) = x_1$, etc. Then we know from the above supposition that the image of $x_n$ in $(C^{(p^n)}/\mathrm{Im}(\lambda^{(p^n)}))(L)$ is a prime to $p$ torsion point for all $n \geq 0$. In particular, the order $m$ of the image of $x_n$ in $(C^{(p^n)}/\mathrm{Im}(\lambda^{(p^n)}))(L)$ is independent of $n$, because the degree of the Verschiebung is always a power of $p$. Let $m$ be the order of $x_0$ (and hence of all the $x_n$). Then $m \cdot x_n \in \mathrm{Im}(\lambda^{(p^n)})(L)$ for all $n$ and thus $m \cdot x_0$ is indefinitely Verschiebung divisible in $\mathrm{Im}(\lambda)(L)$ (because the Verschiebung morphism commutes with morphisms of commutative group schemes). It now suffices to prove that $m \cdot x_0$ is of finite and prime to $p$ order in $\mathrm{Im}(\lambda)(L)$. Hence, we may and do assume that the morphism $\lambda : \mathrm{Tr}_{L|l_0}(C) \to C$ is a surjection.

Now $\lambda$ is also finite and purely inseparable by [Conrad 2006, Theorem 6.12] and it is thus a bijection. We are now given infinitely many $L$-morphisms

$$\ldots (\lambda^{(p^n)})^*(x_n) \to \cdots \to (\lambda^{(p)})^*(x_1) \to \lambda^*(x_0),$$

where $(\lambda^{(p^n)})^*(x_n)$ is the base change by $\lambda^{(p^n)}$ of $x_n$ viewed as a closed subscheme of $C^{(p^n)}$. The $L$-scheme $(\lambda^{(p^n)})^*(x_n)$ is a torsor under the group scheme $(\ker\lambda)^{(p^n)} \simeq \ker\lambda^{(p^n)}$ and according to Lemma 9.3, there is a finite extension $L'$, which splits all the $(\lambda^{(p^n)})^*(x_n)$. We thus obtain an indefinitely Verschiebung divisible point $x_0'$ in $\mathrm{Tr}_{L|l_0}(C)(L')$, whose image in $C(L')$ is $x_0$. Now $\mathrm{Tr}_{L|l_0}(C)_{L'}$ is by definition the base change to $L'$ of an abelian variety over $l_0$; so we are reduced to showing Theorem 1.4 for abelian varieties $C$ that arise by base-change from $l_0$. Lemma 9.4 below thus concludes the proof. □

**Lemma 9.4.** *We have* $\mathrm{IVD}(C, L) \subseteq \mathrm{Tor}^p(C(L))$ *if* $C \simeq C_0 \times_{l_0} L$, *where* $C_0$ *is an abelian variety over* $l_0$.

*Proof of Lemma 9.4.* By [Esnault and Langer 2013, Theorem 6.2 and afterwards] there is an $m \geq 1$ so that $m \cdot x_0 \in C_0(l_0)$. Since $l_0$ is algebraically closed, this implies that $x_0 \in C_0(l_0)$, concluding the proof. □

*Proof of Theorem 1.4.* We begin with a couple of reductions.

(1) *We may assume in the statement of Theorem 1.4 that* $L$ *is the function field of a smooth and proper curve* $B$ *over* $l_0$.

Using Proposition 9.2 and Proposition 9.1 (d), we see that when carrying out reduction (1), we may assume that $\mathrm{Tr}_{L|l_0}(C) = 0$. Reduction (1) now follows from a standard spreading out argument together with Proposition C.1 in the Appendix. Here one could probably appeal instead to Hilbert's irreducibility theorem (as in [Lang 1983, Chapter 9, Corollary 6.3]) but for lack of an adequate reference in the case of function fields, we prefer to use Proposition C.1.

(2) *We may assume in the statement of Theorem 1.4 that* $\dim(\mathrm{Tr}_{\bar{L}|l_0}(C_{\bar{L}})) = \dim(\mathrm{Tr}_{L|l_0}(C))$.

To see this, suppose for the space of this paragraph that we know that Theorem 1.4 is true in general under restrictions (1) and (2). Let $L'|L$ be a finite extension such that $\dim(\mathrm{Tr}_{L'|l_0}(C_{L'}))$ is maximal among all finite extensions of $L$. In particular we then have $\dim(\mathrm{Tr}_{L'|l_0}(C_{L'})) = \dim(\mathrm{Tr}_{\bar{L}|l_0}(C_{\bar{L}}))$. According to Proposition 9.1 (c), we may assume that $L'|L$ is separable. Replacing $L'$ by the Galois closure of $L'$ over $L$, we may even suppose that $L'|L$ is Galois. Let $y_0 \in C(L)$ be an indefinitely Verschiebung divisible element. Suppose $y_0 \neq 0$. Applying our assumptions to $C_{L'}$ and to the normalisation $B'$ of $B$ in $L'$, we see that the image of $y_0$ in $C_{L'}(L')$ is indefinitely Verschiebung divisible. Thus for some integer $m_{y_0}$, which is prime to $p$, the element $m_{y_0} \cdot y_0$ is divisible in the group $C_{L'}(L')$. Now there is a natural group homomorphism $u : C_{L'}(L') \to C(L)$ (the trace) given by the formula

$$u(z) = \sum_{\sigma \in \mathrm{Gal}(L'|L)} \sigma(z).$$

Hence $m_{y_0} \cdot u(y_0) = m_{y_0} \cdot [L' : L] \cdot y_0$ is divisible in the group $C(L)$ and hence

$$m_{y_0} \cdot [L' : L] \cdot y_0 \in \mathrm{Tr}_{L|l_0}(C)(l_0).$$

Now if the order of the image of $y_0$ in $C(L)/\mathrm{Tr}_{L|l_0}(C(l_0))$ is prime to $p$ then we are done. Otherwise, we may (and do) replace $y_0$ by a multiple such that the image in $C(L)/\mathrm{Tr}_{L|l_0}(C(l_0))$ of $y_0$ is a nonzero element of order $p$. In the rest of the argument, we shall derive a contradiction from the existence of this element. Let $i \geq 1$. Let $y_i \in C^{(p^i)}(L)$ be such that $V^{(i)}_{C^{(p^i)}/L}(y_1) = y_0$. The variety

$$(C^{(p^i)})_{L'} = (C_{L'})^{(p^i)} \equiv C^{(p^i)}_{L'}$$

also has the property that $\dim(\mathrm{Tr}_{L'|l_0}(C^{(p^i)}_{L'})) = \dim(\mathrm{Tr}_{\bar{L}|l_0}(C^{(p^i)}_{\bar{L}}))$ since $C^{(p^i)}$ is isogenous to $C$ over $L$. Therefore, repeating the above reasoning, there is an integer $m_{y_i}$, which is prime to $p$, such that $m_{y_i} \cdot [L' : L] \cdot y_i \in \mathrm{Tr}_{L|l_0}(C^{(p^i)}(L)$. Now according to Proposition 9.1 (c), the natural morphism

$\mathrm{Tr}_{L|l_0}(C)_L^{(p^i)} \to C^{(p^i)}$ obtained by base change under $F_C^{\circ i}$ from the morphism $\mathrm{Tr}_{L|l_0}(C)_L \to C$ makes $\mathrm{Tr}_{L|l_0}(C)^{(p^i)}$ into the trace of $C^{(p^i)}$. Thus the map $V^{(i)}_{C^{(p^i)}/L}(\bar{L})$ induces a surjective map

$$C^{(p^i)}(\bar{L})/\mathrm{Tr}_{L|l_0}(C)^{(p^i)}(l_0) \to C(\bar{L})/\mathrm{Tr}_{L|l_0}(C)(l_0)$$

and the map $F^{(i)}_{C/L}(\bar{L})$ induces a bijective map

$$C(\bar{L})/\mathrm{Tr}_{L|l_0}(C)(l_0) \to C^{(p^i)}(\bar{L})/\mathrm{Tr}_{L|l_0}(C)^{(p^i)}(l_0).$$

Since $V^{(i)}_{C^{(p^i)}/L}(\bar{L}) \circ F^{(i)}_{C/L}(\bar{L}) = p^i$, we see that the order of $y_i$ in

$$C^{(p^i)}(L)/\mathrm{Tr}_{L|l_0}(C^{(p^i)})(l_0) \subseteq C^{(p^i)}(\bar{L})/\mathrm{Tr}_{L|l_0}(C^{(p^i)})(l_0)$$

is $p^{i+1}$. This is a contradiction if $i$ is chosen large enough so that $p^i$ is not a divisor of $[L' : L]$. We conclude that the order of the image of $y_0$ in $C(L)/\mathrm{Tr}_{L|l_0}(C)(l_0)$ is prime to $p$ and this concludes reduction step (2).

We now assume that we are given an abelian variety $C$ over $L$ and that $C$ satisfies the assumptions of Theorem 1.4 as well as (1) and (2).

Let as before $\lambda : \mathrm{Tr}_{L|l_0}(C)_L \to C$ be the canonical morphism. According to Proposition 9.2, it will be sufficient to prove that $\mathrm{IVD}(C/\mathrm{Im}(\lambda), L) \subseteq \mathrm{Tor}^p((C/\mathrm{Im}(\lambda))(L))$. By Proposition 9.1 (d), we have $\mathrm{Tr}_{L|l_0}(C/\mathrm{Im}(\lambda)) = 0$ and since we work under supplementary assumption (2), we even have $\mathrm{Tr}_{\bar{L}|l_0}(C/\mathrm{Im}(\lambda)) = 0$. Thus we may replace $C$ by $C/\mathrm{Im}(\lambda)$ and assume from now on that $\mathrm{Tr}_{\bar{L}|l_0}(C) = 0$. Finally, since we have $\mathrm{Tr}_{\bar{L}|l_0}(C) = 0$, we may replace without restriction of generality replace $L$ by a finite extension $L'$ and $B$ by its normalisation $B'$ in $L'$. We may thus assume that there is an integer $m \geq 3$, with $(m, p) = 1$ and such that $C[m] \simeq (\mathbb{Z}/m\mathbb{Z})^{2\dim(C)}$ and $C^\vee[m] \simeq (\mathbb{Z}/m\mathbb{Z})^{2\dim(C^\vee)}$.

By a theorem of Raynaud (see [Abbes 2000, Proposition 5.10]), the connected component of the Néron model of $C$ will then be a semiabelian scheme. We call it $\mathcal{C}$.

Now suppose as in the statement of Conjecture 1.3 that we are given points $x_\ell \in C^{(p^\ell)}(L)$ and suppose that for all $\ell \geqslant 1$, we have $V_{C^{(p^\ell)}/L}(x_\ell) = x_{\ell-1}$. We want to show that $x_0 \in \mathrm{Tor}^p(C(L))$.

By Lemma B.2 and the discussion preceding it we have a canonical map

$$\alpha : C^{(p)}(L) \to \mathrm{Hom}_B(\omega_{\mathcal{C}^{(p)}}, \Omega_{B/l_0}(E)) \tag{7}$$

such that $\alpha(x) = 0$ if and only if $x \in F_{C/L}(C(L))$. Here $E = E(\mathcal{C})$ is the reduced divisor, which is the union of the closed point $b \in B$ such that $\mathcal{C}_b$ is not proper over $\kappa(b)$. Note that we have $E(\mathcal{C}) = E(\mathcal{C}^{(p)}) = E(\mathcal{C}^{(p^2)}) = \cdots$. The map $\alpha$ is naturally compatible with isogenies (we skip the verification) and so there is an infinite commutative diagram

$$\begin{array}{ccc}
C^{(p)}(L) & \longrightarrow & \mathrm{Hom}_B(\omega_{\mathcal{C}^{(p)}}, \Omega_{B/l_0}(E)) \\
{\scriptstyle V_{C^{(p^2)}/L}}\uparrow & & \uparrow{\scriptstyle V^*_{\mathcal{C}^{(p^2)}/B}} \\
C^{(p^2)}(L) & \longrightarrow & \mathrm{Hom}_B(\omega_{\mathcal{C}^{(p^2)}}, \Omega_{B/l_0}(E)) \\
\uparrow & & \uparrow \\
\vdots & \longrightarrow & \vdots
\end{array} \tag{8}$$

Remember that we have

$$\omega_{\mathcal{C}^{(p^n)}} \simeq F_B^{\circ n,*}(\omega_{\mathcal{C}}).$$

Now choose $n_1 \geq 1$ so that

- $\omega_{\mathcal{C}^{(p^{n_1})}}$ has a Frobenius semistable HN filtration;
- $(\omega_{\mathcal{C}^{(p^{n_1})}})_{=0} \simeq (\omega_{\mathcal{C}^{(p^{n_1})}})_{=0,\mathrm{binf}} \oplus (\omega_{\mathcal{C}^{(p^{n_1})}})_{=0,\mu}$ splits into a biinfinitesimal and a multiplicative commutative coLie-algebra (see Lemmata 4.4 and 4.7).

Note that if some $n_1 \geq 1$ has the two above properties, than any higher $n_1$ will as well (by definition for the first property and tautologically for the second one).

Choose $n_2 > n_1$ so that

(I) the image of the map

$$V^{(n_2-n_1),*}_{\mathcal{C}^{(p^{n_2})}/B} : \omega_{\mathcal{C}^{(p^{n_1})}} \to \omega_{\mathcal{C}^{(p^{n_2})}}$$

lies in $(\omega_{\mathcal{C}^{(p^{n_2})}})_{\geq 0} \simeq F_B^{\circ(n_2-n_1),*}((\omega_{\mathcal{C}^{(p^{n_1})}})_{\geq 0})$;

(II) the image of the map of coLie algebras

$$V^{(n_2-n_1),*}_{\mathcal{C}^{(p^{n_2})}/B} : (\omega_{\mathcal{C}^{(p^{n_1})}})_{=0} \to F_B^{\circ(n_2-n_1),*}((\omega_{\mathcal{C}^{(p^{n_1})}})_{=0}) = (\omega_{\mathcal{C}^{(p^{n_2})}})_{=0}$$

is $F_B^{\circ(n_2-n_1),*}((\omega_{\mathcal{C}^{(p^{n_1})}})_{=0,\mu})$. Note that this is possible because the biinfinitesimal part of $(\omega_{\mathcal{C}^{(p^{n_1})}})_{=0}$ will be sent to 0 by sufficiently many composed Verschiebung morphisms (by definition).

Note that under (I) for any $n_3 > n_2$ the image of the map

$$V^{(n_3-n_2),*}_{\mathcal{C}^{(p^{n_3})}/B} : (\omega_{\mathcal{C}^{(p^{n_2})}})_{\geq 0} \to \omega_{\mathcal{C}^{(p^{n_3})}}$$

and hence of the map

$$V^{(n_3-n_1),*}_{\mathcal{C}^{(p^{n_3})}/B} : \omega_{\mathcal{C}^{(p^{n_1})}} \to \omega_{\mathcal{C}^{(p^{n_3})}}$$

automatically lies in $(\omega_{\mathcal{C}^{(p^{n_3})}})_{\geq 0} \simeq F_B^{\circ(n_3-n_1),*}((\omega_{\mathcal{C}^{(p^{n_1})}})_{\geq 0})$.

Choose $n_3 > n_2$ so that

(III) the map

$$\omega_{\mathcal{C}^{(p^{n_3})}} \to \Omega_{B/l_0}(E)$$

given by $x_{n_3}$ factors through its quotient $(F_B^{\circ n_3,*}(\omega_{\mathcal{C}}))_{\leq 0} \simeq F_B^{\circ(n_3-n_1),*}((\omega_{\mathcal{C}^{(p^{n_1})}})_{\leq 0})$;

(IV) the image of the map

$$V^{(n_3-n_2),*}_{\mathcal{C}^{(p^{n_3})}/B} : F_B^{\circ(n_2-n_1),*}((\omega_{\mathcal{C}^{(p^{n_1})}})_{=0}) \to F_B^{\circ(n_3-n_2),*}((\omega_{\mathcal{C}^{(p^{n_2})}})_{=0})$$

is $F_B^{\circ(n_3-n_2),*}((\omega_{\mathcal{C}^{(p^{n_2})}})_{=0,\mu}) \simeq F_B^{\circ(n_3-n_1),*}((\omega_{\mathcal{C}^{(p^{n_1})}})_{=0,\mu})$.

Now we shall exploit the compatibility between the morphism

$$\omega_{\mathcal{C}^{(p^{n_1})}} \xrightarrow{c(x_{n_1})} \Omega_{B/k}(E)$$

induced by $x_{n_1}$ and the morphism

$$\omega_{\mathcal{C}^{(p^{n_3})}} \xrightarrow{c(x_{n_3})} \Omega_{B/k}(E)$$

induced by $x_{n_3}$. According to the diagram (8), this compatibility gives the equality

$$c(x_{n_3}) \circ V^*_{\mathcal{C}^{(p^{n_3-n_1})}/B} = c(x_{n_1}).$$

In other words the composition of morphisms

$$\omega_{\mathcal{C}^{(p^{n_1})}} \xrightarrow{V^*_{\mathcal{C}^{(p^{n_3-n_1})}/B}} \omega_{\mathcal{C}^{(p^{n_3})}} \xrightarrow{c(x_{n_3})} \Omega_{B/k}(E)$$

is $c(x_{n_1})$. Furthermore, in view of (I) and (III) the map $c(x_{n_1})$ factors as

$$\omega_{\mathcal{C}^{(p^{n_1})}} \xrightarrow{V^{(n_3-n_1),*}_{\mathcal{C}^{(p^{n_3})}/B}} F_B^{\circ(n_3-n_1),*}((\omega_{\mathcal{C}^{(p^{n_1})}})_{\geq 0}) \to F_B^{\circ(n_3-n_1),*}((\omega_{\mathcal{C}^{(p^{n_1})}})_{=0}) \to F_B^{\circ(n_3-n_1),*}((\omega_{\mathcal{C}^{(p^{n_1})}})_{\leq 0}) \to \Omega_{B/k}(E)$$

and by (I) the map

$$\omega_{\mathcal{C}^{(p^{n_1})}} \xrightarrow{V^{(n_3-n_1),*}_{\mathcal{C}^{(p^{n_3})}/B}} F_B^{\circ(n_3-n_1),*}((\omega_{\mathcal{C}^{(p^{n_1})}})_{=0})$$

factors as

$$\omega_{\mathcal{C}^{(p^{n_1})}} \xrightarrow{V^{(n_2-n_1),*}_{\mathcal{C}^{(p^{n_1})}/B}} F_B^{\circ(n_2-n_1),*}((\omega_{\mathcal{C}^{(p^{n_1})}})_{\geq 0}) \to F_B^{\circ(n_2-n_1),*}((\omega_{\mathcal{C}^{(p^{n_1})}})_{=0}) \xrightarrow{V^{(n_3-n_2),*}_{\mathcal{C}^{(p^{n_1})}/B}} F_B^{\circ(n_3-n_1),*}((\omega_{\mathcal{C}^{(p^{n_1})}})_{=0})$$

and thus by (IV) and (II) the image of this last map is precisely $F_B^{\circ(n_3-n_1),*}((\omega_{\mathcal{C}^{(p^{n_1})}})_{=0,\mu})$.

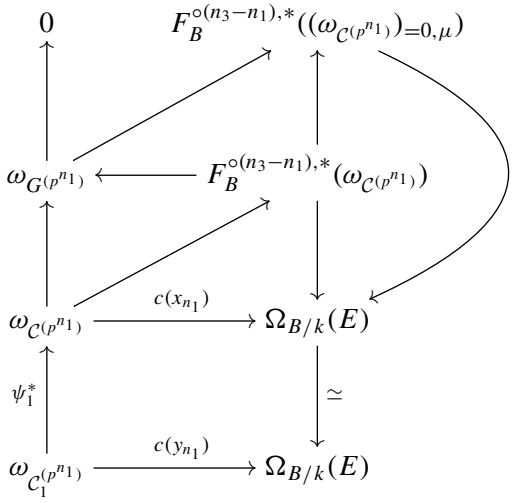
We have thus constructed a multiplicative quotient of the $p$-coLie algebra $\omega_{\mathcal{C}^{(p^{n_1})}}$. On the other hand the $p$-coLie algebra $\omega_{\mathcal{C}^{(p^{n_1})}}$ is the $p$-coLie algebra of the finite flat group scheme $\ker F_{\mathcal{C}^{(p^{n_1})}/B}$. By the equivalence of categories recalled in Section 4B, this quotient corresponds to a multiplicative subgroup scheme of $\ker F_{\mathcal{C}^{(p^{n_1})}/B}$. By Lemma 4.8, this subgroup scheme embeds in the canonical largest multiplicative subgroup scheme $(\ker F_{\mathcal{C}^{(p^{n_1})}/B})_\mu$ of $\ker F_{\mathcal{C}^{(p^{n_1})}/B}$ (in fact, it coincides with it, but we shall not need this). Finally note that

$$(\ker F_{\mathcal{C}^{(p^{n_1})}/B})_\mu \simeq ((\ker F_{\mathcal{C}/B})_\mu)^{(p^{n_1})},$$

by the last part of Lemma 4.8.

Let $G := (\ker F_{\mathcal{C}/B})_\mu$. Note that $G = G_{\mathcal{C}}$ in the notation of Theorem 2.1. Now consider the quotient $\mathcal{C}_1 := \mathcal{C}/G$ (which is a semiabelian scheme by 4.10) and let $\psi_1 : \mathcal{C} \to \mathcal{C}_1$ be the quotient morphism. The

point $x_{n_1}$ and its image $y_{n_1}$ in $\mathcal{C}_1(L)$ give a commutative diagram




where the left column is an exact sequence and $c(y_{n_1})$ is the morphism induced by $y_{n_1}$.

*Thus $c(y_{n_1})$ vanishes.* In particular, $y_{n_1}$ lies in the image of $F_{\mathcal{C}_1^{(p^{n_1-1})}/B}(\mathcal{C}_1^{(p^{n_1-1})}(L))$. Using the fact that

$$[p]_{\mathcal{C}_1^{(p^{n_1-1})}} = V_{\mathcal{C}_1^{(p^{n_1})}/B} \circ F_{\mathcal{C}_1^{(p^{n_1-1})}/B},$$

we conclude that $y_{n_1-1}$ has a $p$-th root in $\mathcal{C}_1^{(p^{n_1-1})}(L)$. Hence $y_0$ also has a $p$-th root in $\mathcal{C}_1(L)$. Now since $G$ is independent of $x_0$, we conclude that the image of any indefinitely Verschiebung divisible point of $C(L)$ in $\mathcal{C}_1(L)$ has a $p$-th root. Since $G$ is compatible with twists, we also see that for any $n \geq 0$ the image of any indefinitely Verschiebung divisible point of $C^{(p^n)}(L)$ in ${\mathcal{C}_1}^{(p^n)}(L)$ has a $p$-th root. From this, by an elementary combinatorial consideration, we see that the image of any indefinitely Verschiebung divisible point of $C(L)$ in $\mathcal{C}_1(L)$ has a $p$-th root, which is indefinitely Verschiebung divisible.

By the discussion above, *the image of* $\mathrm{IVD}(C)$ *in* $\mathcal{C}_1(L)$ *lies in* $p \cdot \mathrm{IVD}(\mathcal{C}_{1,L})$. This is the crucial fact that the rest of the proof will exploit.

Let $\mathcal{C}_1 := \mathcal{C}/G_{\mathcal{C}}$, $\mathcal{C}_2/G_{\mathcal{C}_1}, \ldots$ be the sequence of smooth commutative group schemes obtained by successively quotienting by the canonical subgroup schemes described in Theorem 2.1. Note that all the $\mathcal{C}_i$ are semiabelian by Lemma 4.10. We shall denote by $\psi_i$ the morphism $\mathcal{C} \to \mathcal{C}_i$ obtained by composition. We write $C_i := \mathcal{C}_{i,L}$ for convenience.

Let $m_{00}$ be an integer such that $m_{00} \cdot x_0 =: v_0$ extends to an element $\tilde{v}_0$ of $\mathcal{C}(B)$.

Now let $D_0$ be a line bundle on $C$. We suppose that $[-1]_C^*(D_0) \simeq D_0$ (i.e., $D_0$ is symmetric), where $[-1]_C$ is the inversion morphism given by the group scheme structure of $C$ over $B$. We also suppose that $D_0$ is a relatively ample line bundle. If $x \in \mathcal{C}(B)$, write $\tau_x : \mathcal{C} \to \mathcal{C}$ for the translation by $x$ morphism. We use the same notation for $x \in C(L)$.

Now consider the isogeny $\phi_{D_0} : C \to C^\vee$ from $C$ to its dual abelian variety, which is induced by $D_0$ (this is the polarisation induced by $D_0$). Since $v_0 \in \mathrm{IVD}(C)$, we also have $\phi_{D_0}(v_0) \in \mathrm{IVD}(C^\vee)$, since

relative Frobenius morphisms are naturally compatible with morphisms of abelian varieties. The point $\phi_{D_0}(v_0)$ corresponds to the line bundle

$$M = \tau_{v_0}^*(D_0) \otimes D_0^\vee$$

on $C$ (see [Mumford 1970, III.13]). Since the morphism dual to the Verschiebung morphism is the relative Frobenius morphism (this is very often the definition of the Verschiebung), we see that the fact that $\phi_{D_0}(v_0) \in \mathrm{IVD}(C^\vee)$ translates to the fact that there exist line bundles $M_i$ on $C^{(p^i)}$ for all $i \geq 1$, such that

$$F_{C/L}^*(M_1) \simeq M, \quad F_{C^{(p)}/L}^*(M_2) \simeq M_1, \quad F_{C^{(p^2)}/L}^*(M_3) \simeq M_2, \quad \dots$$

Since $\psi_i$ factors by construction through $F_{C^{(p^{i-1})}/L} \circ F_{C^{(p^{i-2})}/L} \circ \cdots \circ F_{C/L}$, we see that for each $i \geq 1$, there is a line bundle $J_i$ on $C_i$ such that $\psi_{i,L}^*(J_i) \simeq M$.

Now recall that $D_0$ extends uniquely (up to isomorphism) to a line bundle $\mathcal{D}_0$ on $\mathcal{C}$, if we require $D_0$ to be trivial along the unit section of $\mathcal{C}$ (see [Moret-Bailly 1985, Proposition 2.6, p. 21]). Similarly the line bundle $M$ extends uniquely (up to isomorphism) to a line bundle $\mathcal{M}$ on $\mathcal{C}$ with the same property. We shall write $\mathcal{J}_i$ for the line bundle similarly associated with $J_i$ on $\mathcal{C}_i$. Notice that by unicity, we have $\psi_i^*(\mathcal{J}_i) \simeq \mathcal{M}$.

We shall now make a height computation. We shall need:

**Lemma 9.5.** *Let $\mathcal{W}$ be a line bundle on $\mathcal{C}$, which is trivial when restricted to the unit section and such that $W_L$ is algebraically equivalent to 0. Let $x \in \mathcal{C}(B)$. Then $\deg(x^*(\mathcal{W}))$ is the Néron–Tate height pairing of $x_L \in C(L)$ and $\mathcal{W}_L$.*

*Proof.* This follows from [Moret-Bailly 1985, III.3.2 and 3.3] and the definition of polarisations. □

**Proposition 9.6.** (a) *There exists a constant $m_0 \in \mathbb{N}^*$ and an infinite set $I_0 \subseteq \mathbb{N}^*$ such that for any $i \in I_0$ and any $P \in C_i(L)$, the element $m_0 \cdot P$ extends to an element of $\mathcal{C}_i(B)$.*

(b) *There is a constant $c_0 \in \mathbb{N}^*$ and an infinite set $I_0 \subseteq \mathbb{N}^*$ such that for any $i \in I_0$ and any $P \in \mathrm{Tor}(C_i(L))$ we have $c_0 \cdot P = 0$.*

We shall prove this proposition later, using Proposition A.2 in the Appendix.

Let $i \in I_0$. For the next computation, recall that $\psi_{i,L}(v_0)$ is divisible by $p^i$ in $C_i(L)$. Let $z_i$ be an element of $C_i(L)$ such that $p^i \cdot z_i = \psi_{i,L}(v_0)$. According to Proposition 9.6 (a), $m_0 \cdot z_i$ extends to an element $u_i$ of $\mathcal{C}_i(B)$. By construction, we have $p^i \cdot u_i = m_0 \cdot \psi_i(\tilde{v}_0)$. We compute

$$\begin{aligned}\deg(([m_0](\tilde{v}_0))^*(\mathcal{M})) &= \deg(([m_0](\tilde{v}_0))^*(\psi_i^*(J_i))) = \deg(([m_0](\psi_i(\tilde{v}_0)))^*(J_i)) \\ &= \deg(([p^i](u_i))^*(J_i)) = \deg(u_i^*([p^i]^*(J_i))) \\ &= \deg(u_i^*(J_i^{\otimes p^i})) = p^i \cdot \deg(u_i^*(J_i)).\end{aligned}$$

Here $[m_0]$ refers to the multiplication by $m_0$ morphism (in particular $[m_0](\tilde{v}_0) = m_0 \cdot \tilde{v}_0$). Suppose for contradiction that $\deg(([m_0](\tilde{v}_0))^*(\mathcal{M})) \neq 0$. If we choose $i$ large enough so that $p^i$ is not a divisor of $\deg(([m_0](\tilde{v}_0))^*(\mathcal{M}))$ then we get a contradiction. Thus $\deg(([m_0](\tilde{v}_0))^*(\mathcal{M})) = 0$. We may also compute

$$\deg(([m_0](\tilde{v}_0))^*(\mathcal{M})) = \deg(\tilde{v}_0^*([m_0]^*(\mathcal{M}))) = \deg(\tilde{v}_0^*(\mathcal{M}^{\otimes m_0})) = m_0 \cdot \deg(\tilde{v}_0^*(\mathcal{M})).$$

In particular, by Lemma 9.5, the Néron–Tate height pairing of $v_0$ and $M$ vanishes. Now notice that $M$ is by definition the image of $v_0$ under the polarisation induced by the symmetric ample line bundle $D_0$. Hence the Néron–Tate pairing of $v_0$ and $M$ is twice the Néron–Tate height of $v_0$ with respect to the polarisation induced by $D_0$. In particular, the Néron–Tate height of $v_0$ with respect to $D_0$ vanishes. By a theorem of Lang (see [Conrad 2006, Theorem 9.15]) we conclude that the image of $v_0$ in $C(L)$ is an element of finite order. Thus the image of $x_0$ in $C(L)$ is also an element of finite order.

Now we show that $x_0 \in \mathrm{Tor}^p(C(L))$. For contradiction, suppose that $x_0 \notin \mathrm{Tor}^p(C(L))$. We thus may (and do) replace $x_0$ by one of its multiples and suppose that $p \cdot x_0 = 0$ and $x_0 \neq 0$. We know that $\psi_{i,L}(x_0)$ is divisible by $p^i$ in $C_i(L)$ and since $\psi_{i,L}$ is injective we conclude that there is an element of order $p^{i+1}$ in $C_i(L)$ for all $i \geq 1$. This contradicts Proposition 9.6 (b) so we are done. □

*Proof of Proposition 9.6.* We need some preliminaries on moduli spaces of abelian varieties. Let $n \geq 3$ with $(n, p) = 1$) and $g \geq 1$. We shall choose particular values for $g$ and $n$ later.

Let $\boldsymbol{A}_{g,n}$ be the functor from the category of locally noetherian $\mathbb{F}_p$-schemes to the category of sets, such that

$$\boldsymbol{A}_{g,n}(B) = \Big\{\text{isomorphism classes of the following objects: principally polarised abelian schemes over } B \text{ endowed with a symplectic isomorphism } (\mathbb{Z}/n\mathbb{Z})_B^{2g} \simeq \mathcal{A}[n]\Big\}$$

D. Mumford proved (see [Mumford et al. 1994]) that the functor $\boldsymbol{A}_{g,n}$ is representable by a scheme, which is separated and of finite type over $\mathbb{F}_p$. We shall also denote this scheme by $\boldsymbol{A}_{g,n}$.

Furthermore, C. Chai and G. Faltings [1990, V, 2., Theorem 2.5] proved that there exists

- a scheme $\bar{\boldsymbol{A}}_{g,n}$ (resp. $\boldsymbol{A}^*_{g,n}$), which is proper over $\mathbb{F}_p$;
- an open immersion $\boldsymbol{A}_{g,n} \hookrightarrow \bar{\boldsymbol{A}}_{g,n}$ (resp. an open immersion $\boldsymbol{A}_{g,n} \hookrightarrow \boldsymbol{A}^*_{g,n}$);
- a semiabelian scheme $\mathcal{U}$ over $\bar{\boldsymbol{A}}_{g,n}$, such that $\mathcal{U}_{\boldsymbol{A}_{g,n}}$ is isomorphic to the universal abelian scheme over $\boldsymbol{A}_{g,n}$;
- a morphism $\bar{\pi} : \bar{\boldsymbol{A}}_{g,n} \to \boldsymbol{A}^*_{g,n}$ compatible with the above open immersions of $\boldsymbol{A}_{g,n}$;
- a line bundle $\omega^0$ on $\boldsymbol{A}^*_{g,n}$, which is ample and such that $\bar{\pi}^*(\omega^0) = \omega_{\mathcal{U}/\bar{\boldsymbol{A}}_{g,n}}$.

Write $Z := B \times_{l_0} \boldsymbol{A}^*_{g,n,l_0}$. Recall that the Hilbert scheme $\mathrm{Hilb}(Z/l_0)$ is a scheme, representing the functor

$$T \mapsto \{\text{closed subschemes of } Z_T, \text{ which are proper and flat over } T\}$$

from the category of locally noetherian scheme $T$ over $l_0$ to the category of sets. It is locally of finite type over $l_0$ (see [Grothendieck 1966]).

Let $\Phi \in \mathbb{Q}[\lambda]$ be a polynomial with rational coefficients and $L_0/Z$ an ample line bundle. By definition, the $l_0$-scheme $\mathrm{Hilb}_\Phi(Z/l_0)$ represents the functor

$$T \mapsto \Big\{\text{closed subschemes } W \text{ of } Z_T, \text{ which are proper and flat over } T \text{ and such that } \chi(W_t, L_{0,W_t}^{\otimes\lambda}) = \Phi(\lambda) \text{ for all } \lambda \in \mathbb{N} \text{ and all } t \in T\Big\}$$

from the category of locally noetherian scheme $T$ over $l_0$ to the category of sets.

Here $W_t$ is the fibre at $t \in T$ of the morphism $W \to T$ and $L_{0,W_t}$ is the pull-back of $L$ to $W_t$ by the natural morphism $W_t \to Z$. The symbol $\chi(\cdot)$ refers to the Euler characteristic. By definition

$$\chi(W_t, L_{W_t}^{\otimes\lambda}) = \sum_{r\geqslant 0}(-1)^r \dim_{\kappa(t)} H^r(W_t, L_{W_t}^{\otimes\lambda}).$$

(this is called the Hilbert polynomial of $W_t$ with respect to $L_{W_t}$). It is shown in [Grothendieck 1966], that $\mathrm{Hilb}_\Phi(Z/l_0)$ is projective over $l_0$ (as a consequence of the projectivity of $Z$). Notice that by construction, we have a disjoint union

$$\mathrm{Hilb}(Z/l_0) = \coprod_{\Phi\in\mathbb{Q}[\lambda]} \mathrm{Hilb}_\Phi(Z/l_0)$$

Finally, it is shown in [Fantechi et al. 2005, part II, 5.23] that the functor $\mathrm{Mor}_{l_0}(B, \boldsymbol{A}^*_{g,n})$ from locally noetherian $l_0$-schemes $T$ to the category of sets, such that

$$\mathrm{Mor}_{l_0}(B, \boldsymbol{A}^*_{g,n,l_0})(T) = \{T\text{-morphisms from } B_T \text{ to } \boldsymbol{A}^*_{g,n,T}\}$$

is representable by an open subscheme of $\mathrm{Hilb}(Z/l_0)$. More precisely, the natural transformation of functors

$$T\text{-morphism } f \text{ from } B_T \text{ to } \boldsymbol{A}^*_{g,n,T} \mapsto \text{graph of } f$$

is represented by an open immersion

$$\mathrm{Mor}_{l_0}(B, \boldsymbol{A}^*_{g,n}) \hookrightarrow \mathrm{Hilb}(B \times_{l_0} \boldsymbol{A}^*_{g,n,l_0}/l_0).$$

Let now $D$ be an ample line bundle on $B$. We choose $L_0$ to be the line bundle $D\boxtimes\omega_{l_0}^0$ on $Z = B\times_{l_0}\boldsymbol{A}^*_{g,n,l_0}$.

Recall that the Hodge bundles of the $\mathcal{C}_i$ all have the same degree by Lemma 4.12. Let $d_0 := \deg(\omega_{\mathcal{C}/B})$ be this common degree. Our aim is to use this to show that all the $C_i$ embed in a bounded family of abelian varieties and apply Proposition A.2.

Notice that $C_i[m] \simeq (\mathbb{Z}/m\mathbb{Z})^{2\dim(C_i)}$ and $C_i^\vee[m] \simeq (\mathbb{Z}/m\mathbb{Z})^{2\dim(C_i^\vee)}$. Indeed, since $\psi_{i,L}$ is purely inseparable, it induces an isomorphism $C[m] \to C_i[m]$ and thus $C_i[m] \simeq (\mathbb{Z}/m\mathbb{Z})^{2\dim(C_i)}$ by (IV) above. For the isomorphism $C_i^\vee[m] \simeq (\mathbb{Z}/m\mathbb{Z})^{2\dim(C_i^\vee)}$, notice that the dual morphism $\psi_{i,L}^\vee : C_i^\vee \to C^\vee$ is separable (because its kernel is the Cartier dual of a multiplicative group scheme) and of order a power of $p$. Hence, since $(p, m) = 1$ it also induces an isomorphism $C_i^\vee[m] \to C^\vee[m]$ (we leave the details to the reader).

Now let $E_i := (C_i \times_L C_i^\vee)^4$. By Zarhin's trick (see [Moret-Bailly 1985, IX.1.1]) $E_i$ carries a principal polarisation. Furthermore, by the last paragraph, we also have $E_i[m] \simeq (\mathbb{Z}/m\mathbb{Z})^{2\dim(E_i)}$. Notice also that the identity component of the Néron model of $C_i$ is semiabelian, since $\mathcal{C}_i$ is semiabelian. Hence the identity component of the Néron model of $C_i^\vee$ is also semiabelian, since $C_i^\vee$ is isogenous to $C_i$ (see [Abbes 2000, Proposition 5.8 (4)] for a neat presentation). Since the formation of the Néron model is compatible with products, we conclude that the identity component $\mathcal{E}_i$ of the Néron model of $E_i$ is also semiabelian. We also see $\mathcal{E}_i|_{B\setminus E(\mathcal{C})}$ is an abelian scheme over $B\setminus E(\mathcal{C})$ (where $E(\mathcal{C})$ is as in (7)). Finally, we have $\deg(\omega_{\mathcal{E}_i/B}) = 8\cdot d_0$ by [Faltings and Chai 1990, V.3, Lemma 3.4, p. 166].

Let now $g = \dim(E_i) = 8 \cdot \dim(C)$ and $n = m$. By definition, $E_i$ is associated with an $l_0$-morphism $\operatorname{Spec} L \to \boldsymbol{A}_{g,n,l_0}$. By the valuative criterion of properness, this morphism extends to a morphism $\phi_i : B \to \boldsymbol{A}^*_{g,n,l_0}$ (resp. to a morphism $\bar{\phi}_i : B \to \bar{\boldsymbol{A}}_{g,n,l_0}$). By unicity, we have $\bar{\pi} \circ \bar{\phi}_i = \phi_i$. Thus, since semiabelian extensions are unique (see [Raynaud 1970, IX, Corollary 1.4, p. 130]), we have $\phi_i^*(\omega^0_{l_0}) \simeq \omega_{\mathcal{E}_i/B}$. The morphism $\phi_i$ is by definition associated with an element of $\operatorname{Mor}_{l_0}(B, \boldsymbol{A}^*_{g,n,l_0})(l_0)$. We can now compute the Hilbert polynomial of the graph $\Gamma_{\phi_i}$ of $\phi_i$ with respect to the line bundle $L_0$:

$$\begin{aligned}
\chi(\Gamma_{\phi_i}, L_0^{\otimes\lambda}) =: Q(\lambda) &= \chi(B, (D \otimes \phi_i^*(\omega^0_{l_0}))^{\otimes\lambda}) \\
&= \deg_B((D \otimes \phi_i^*(\omega^0_{l_0}))^{\otimes\lambda}) + 1 - g(B) \\
&= \lambda \cdot \deg_B(D \otimes \phi_i^*(\omega^0_{l_0})) + 1 - g(B) \\
&= \lambda \cdot \deg_B(D \otimes \omega_{\mathcal{E}_i/B}) + 1 - g(B) \\
&= \lambda \cdot \deg_B(D) + \lambda \cdot \deg_B(\omega_{\mathcal{E}_i/B}) + 1 - g(B) \\
&= \lambda \cdot \deg_B(D) + \lambda \cdot 8 \cdot d_0 + 1 - g(B). \qquad (9)
\end{aligned}$$

Here $g(B)$ is the genus of $B$. The second equality is justified by the Riemann–Roch theorem on $B$. We thus see that the Hilbert polynomial $Q(\lambda)$ of the graph of $\phi_i$ with respect to $L_0$ is $Q(\lambda)$ is independent of $i$. Thus the element of $\operatorname{Mor}_{l_0}(B, \boldsymbol{A}^*_{g,n,l_0})(l_0)$ corresponding to $\mathcal{E}_i$ lies in the scheme

$$\operatorname{Mor}_{l_0}(B, \boldsymbol{A}^*_{g,n,l_0})(l_0) \cap \operatorname{Hilb}_{Q(\lambda)}(B \times_{l_0} \boldsymbol{A}^*_{g,n,l_0}/l_0)$$

which is of finite type over $l_0$ by the above discussion. We now let $Y$ be the Zariski closure in

$$\operatorname{Mor}_{l_0}(B, \boldsymbol{A}^*_{g,n,l_0})(l_0) \cap \operatorname{Hilb}_{Q(\lambda)}(B \times_{l_0} \boldsymbol{A}^*_{g,n,l_0}/l_0)$$

of the set all the elements of $(\operatorname{Mor}_{l_0}(B, \boldsymbol{A}^*_{g,n,l_0})(l_0) \cap \operatorname{Hilb}_{Q(\lambda)}(B \times_{l_0} \boldsymbol{A}^*_{g,n,l_0}/l_0))(l_0)$ which correspond to some $\phi_i$ ($i \geq 0$). Finally we let $H_{00}$ be some irreducible component of $Y$, which meets infinitely many such points. Let $\eta_{00} := \kappa(H_{00})$. By construction, we have an $H_{00}$-morphism

$$B \times_{l_0} H_{00} \to \boldsymbol{A}^*_{g,n,H_{00}}$$

which sends $(B \backslash E(\mathcal{C}))_{\eta_{00}}$ into $\boldsymbol{A}_{g,n,\eta_{00}} \subseteq \boldsymbol{A}^*_{g,n,\eta_{00}}$ (because by construction, $(B \backslash E(\mathcal{C}))_x$ is sent into $\boldsymbol{A}_{g,n,x}$ for a dense sent of points $x \in H_{00}$). Let

$$\gamma_0 : B_{\eta_{00}} \to \boldsymbol{A}^*_{g,n,\eta_{00}}$$

be the induced morphism over $\eta_{00}$. Now recall that there is a proper morphism $\bar{\pi} : \bar{\boldsymbol{A}}_{g,n} \to \boldsymbol{A}^*_{g,n}$. By the valuative criterion of properness, there is a unique $\eta_{00}$-morphism $\gamma : B_{\eta_{00}} \to \bar{\boldsymbol{A}}_{g,n,\eta_{00}}$ such that $\bar{\pi}_{\eta_{00}} \circ \gamma = \gamma_0$. The morphism $\gamma$ extends over an open subset $H_0$ of $H_{00}$, yielding an $H_0$-morphism

$$\widetilde{\gamma} : B \times_{l_0} H_0 \to \bar{\boldsymbol{A}}_{g,n,H_0}.$$

Replacing $H_0$ by one of its open subsets, we may suppose that $H_0$ is normal. Let now $\mathcal{B}_0$ be the base change of $\mathcal{U}$ by $\widetilde{\gamma}$. A theorem of Moret-Bailly [1985, VI.3.1] together with a result of Raynaud [1970, XI.1.4] then shows that $\mathcal{B}_0$ can be endowed with a relatively ample line bundle, which is symmetric and

trivial along the zero section. Let also $t_0 := l_0$, $C := B \times_{l_0} H_0$. If we now apply Proposition A.2 (a) with this choice of $H_0$, $t_0$, $C$ and $\mathcal{B}_0$, we reach the conclusion that there is an infinite set $I_0 \subseteq \mathbb{N}^*$ and a constant $n_0$, such that for $i \in I_0$, and any $P \in E_i(L)$, the element $n_0 \cdot P$ extends to an element of $\mathcal{E}_i(L)$. Since $\mathcal{C}_i$ is a direct factor of $\mathcal{E}_i$, we may replace $E_i$ (resp. $\mathcal{E}_i$) by $C_i$ (resp. $\mathcal{C}_i$) in the last sentence. This proves (a), with $m_0 = n_0$. For (b), note that $\mathrm{Tr}_{L|l_0}(E_i) = 0$ (since $E_i$ is a product of abelian varieties isogenous to $C$) and apply Proposition A.2 (b) to the same situation. □

## Appendix A: Rational points in families

The terminology of this appendix is independent of the terminology of the rest of the article and appendices.

Let $t_0$ be an algebraically closed field. Let $H_0$ be an integral scheme of finite type over $t_0$. Let $\pi : C \to H_0$ be a smooth curve over $H_0$, with geometrically connected fibres. Let $\mathcal{B}_0$ be a semiabelian scheme over $C$. Suppose that there exists a line bundle $L$ on $\mathcal{B}_0$, which is ample relatively to $C$, symmetric and trivial along the zero section. Let $\eta_0 := \kappa(H_0)$ and let $\lambda_0 := \kappa(C)$. Note that $\lambda_0$ lies over $\eta_0$ via $\pi$ and that $\lambda_0$ is also the generic point of $C_{\eta_0}$ viewed as a subset of $C$. We suppose that $\mathcal{B}_{0,\lambda_0}$ is an abelian variety over $\lambda_0$.

In the next proposition, we shall need the following lemma, which is well known from the theory of minimal models of curves.

**Lemma A.1.** *Let $\phi : X \to Y$ be a morphism of smooth varieties over $t_0$. Suppose also that there is a dense open set $Y_1 \subseteq Y$, such that $\phi|_{Y_1} : \phi^{-1}(Y_1) \to Y_1$ is smooth. Denote by $X^{\mathrm{sm}}$ the maximal open subscheme of $X$, such that $\phi|_{X^{\mathrm{sm}}} \to Y$ is smooth.*

*Let $\sigma \in X(Y)$ be a section of $\phi$. Then $\sigma \in X^{\mathrm{sm}}(Y) \subseteq X(Y)$.*

*Proof.* See [Liu 2002, Example 4.3.25]. □

**Proposition A.2.** (a) *There is a natural number $n_0$ and a dense open set $V \subseteq H_0$ with the following properties. For any $x \in V(t_0)$, $\mathcal{B}_{0,\kappa(C_x)}$ is an abelian variety and for any $P_x \in \mathcal{B}_0(\kappa(C_x))$, the point $n_0 \cdot P_x \in \mathcal{B}_0(\kappa(C_x))$ extends to an element of $\mathcal{N}(\mathcal{B}_{0,\kappa(C_x)})^0(C_x)$.*

(b) *Suppose that $C$ is proper over $H_0$. Suppose that there is a set $T_0 \subseteq H_0(t_0)$, which is dense in $H_0$ and such that for any $x \in T_0$ we have $\mathrm{Tr}_{\kappa(C_x)|t_0}(\mathcal{B}_{0,\kappa(C_x)}) = 0$. Then there is a dense open set $V \subseteq H_0$ and a natural number $b_0$ such that for all $x \in V(t_0)$, we have $\#\mathrm{Tor}(\mathcal{B}_0(\kappa(C_x))) \leq b_0$.*

Here $\mathcal{N}(\mathcal{B}_{0,\kappa(C_x)})^0$ is the connected component of the identity of the Néron model $\mathcal{N}(\mathcal{B}_{0,\kappa(C_x)})$ of $\mathcal{B}_{0,\kappa(C_x)}$ over $C_x$.

*Proof.* We start with (a). We shall write $\bar{\eta}_0$ for an algebraic closure of $\eta_0$. Consider the semiabelian scheme $\mathcal{B}_{0,\bar{\eta}_0}$ over $C_{\bar{\eta}_0}$. According to [Künnemann 1998, Theorem 4.2], there is an open immersion

$$\mathcal{B}_{0,\bar{\eta}_0} \hookrightarrow S_1 \tag{10}$$

of $C_{\bar{\eta}_0}$-schemes, with the following properties: $S_1$ is a regular scheme, which is projective over $C_{\bar{\eta}_0}$ and the open immersion $\mathcal{B}_{0,\bar{\eta}_0} \hookrightarrow S_1$ is an isomorphism when restricted to the open subset of $C_{\bar{\eta}_0}$ over

which $\mathcal{B}_{0,\bar{\eta}_0}$ is an abelian scheme. In particular $S_1$ is smooth over $\bar{\eta}_0$, since $\bar{\eta}_0$ is perfect. There is a finite field extension $\eta \to \eta_0$ and a morphism

$$\mathcal{B}_{0,\eta} \to S \tag{11}$$

of $C_\eta$-schemes, which is model of (10). By flat descent, the morphism $\mathcal{B}_{0,\eta} \to S$ is also an open immersion and $S$ is also smooth over $\eta$ and projective over $C_\eta$. Again by flat descent $\mathcal{B}_{0,\eta} \to S$ is an isomorphism when restricted to the open subset of $C_\eta$ over which $\mathcal{B}_{0,\eta}$ is an abelian scheme.

We now let $g : H \to H_0$ be the normalisation of $H_0$ in $\eta$. Slightly abusing notation, we also denote by $\eta$ the generic point of $H$. Note that $g$ is a finite morphism (see, e.g., [EGA IV$_2$ 1965, p. 214–218]). We let $\mathcal{B}$ be the semiabelian scheme on $C_H$ obtained by base change and we let $\lambda$ be the generic point of $C_H$. Again $\lambda$ lies over $\eta$ via the second projection and is also the generic point of the $C_\eta$. By an elementary constructibility argument, there is a nonempty open set $U \subseteq H$ and an open immersion

$$\mathcal{B}_{C_U} \hookrightarrow \widetilde{S}$$

of $C_U$-schemes, where $\widetilde{S}$ is smooth over $U$ and projective over $C_U$. Furthermore, we may assume that there is an open subset $U' \subseteq C_U$, which surjects onto $U$, with the property that $\mathcal{B}_{U'}$ is an abelian scheme over $U'$ and that the induced morphism $\mathcal{B}_{U'} \hookrightarrow \widetilde{S}_{U'}$ is an isomorphism.

Let $N_0$ be the supremum of the set of values of the function, which associates with any $q \in C_U$ the number of geometric irreducible components of the fibre $\widetilde{S}_q$ of $\widetilde{S}$ over $q$. This function is constructible (see [EGA IV$_3$ 1966, p. 82]) and so $N_0$ is finite.

Now let $y \in U(t_0)$. By construction $\mathcal{B}_{C_y}$ is then a generically abelian semiabelian scheme over $C_y$. We have a canonical $C_y$-morphism $f : (\widetilde{S}_{C_y})^{\mathrm{sm}} \to \mathcal{N}(\mathcal{B}_{\kappa(C_y)})$ by the definition of the Néron model. Let $P_y \in \mathcal{B}(\kappa(C_y))$. The section $P_y$ extends uniquely to a element of $(\widetilde{S}_{C_y})^{\mathrm{sm}}(C_y)$ by the valuative criterion of properness and Lemma A.1. It also extends uniquely to an element of $\mathcal{N}(\mathcal{B}_{\kappa(C_y)})(C_y)$ by the definition of the Néron model. By unicity, these two extensions are compatible with the morphism $f$. Let $s \in C_y(t_0)$. Since the number of irreducible components of $(\widetilde{S}_{C_y})^{\mathrm{sm}}_s$ is $\leq N_0$, we see that the images of the multiples $P_y, 2 \cdot P_y, \ldots$ of $P_y$ in $\mathcal{N}(\mathcal{B}_{\kappa(C_y)})(s)$ are contained in at most $N_0$ components of $\mathcal{N}(\mathcal{B}_{\kappa(C_y)})_s$. Hence the order of the image of $P_y$ in the component group of $\mathcal{N}(\mathcal{B}_{\kappa(C_y)})_s$ is $\leq N_0$. Since $s$ was arbitrary, we see that $N_0! \cdot P_y$ extends to an element of $\mathcal{N}(\mathcal{B}_{\kappa(C_y)})^0(C_y)$. Note also (for use in (b) below) that since $\mathcal{B}_{C_y}$ is semiabelian, $\mathcal{N}(\mathcal{B}_{\kappa(C_y)})^0(C_y)$ naturally identifies with $\mathcal{B}_{C_y}$ by the unicity of semiabelian extensions.

Finally let $V$ be the open set $H_0 \backslash g(H \backslash U)$. By construction, we have $g^{-1}(V) \subseteq U$. Thus every point of $V(t_0)$ lifts to a point of $U(t_0)$ (since $g$ is finite) and we see that $V$ has the required properties.

For the proof of (b) we first let $U$ be as in the proof of (a). We let $\underline{\mathrm{Sec}}^0_U(\mathcal{B}_{C_U}/C_U)$ the functor from locally noetherian $U$-schemes $T$ to sets, such that

$$\underline{\mathrm{Sec}}^0_U(\mathcal{B}_{C_U}/C_U)(T) = \{\text{sections } \sigma \text{ of } \mathcal{B}_{C_T} \to C_T \text{ such that } \deg((\sigma^*(L))_{C_t}) = 0 \text{ for all } t \in T\}.$$

As $\mathcal{B}_{C_U}$ is quasiprojective over $U$, this functor is representable by a scheme $\mathrm{Sec}^0_U(\mathcal{B}_{C_U}/C_U)$ of finite type over $U$. See, e.g., [Nitsure 2005, Example before 5.6.3]. See the proof of Proposition 9.6 for a similar

construction. We leave the details to the reader. Now let $x \in g^{-1}(T_0) \cap U$. We have an identification

$$\begin{aligned}\mathrm{Sec}^0_U(\mathcal{B}_{C_U}/C_U)_x(t_0) &= \mathrm{Sec}^0_x(\mathcal{B}_{C_x}/C_x)(t_0)\\ &= \{P \in \mathcal{B}_{C_x}(C_x) \mid \text{the Néron–Tate height of } P \text{ with respect to } L_{\mathcal{B}_{C_x}} \text{ vanishes}\}.\end{aligned}$$

See [Moret-Bailly 1985, III.3.2 and 3.3]. Since $\mathrm{Tr}_{\kappa(C_x)|t_0}(\mathcal{B}_{0,\kappa(C_x)}) = 0$, a theorem of Lang (see [Conrad 2006, Theorem 9.15]) implies that $\mathrm{Sec}^0_U(\mathcal{B}_{C_U}/C_U)_x(t_0)$ consists of torsion sections. Furthermore, by the Lang–Néron theorem, $\mathrm{Sec}^0_U(\mathcal{B}_{C_U}/C_U)_x(t_0)$ is finite. Hence $\mathrm{Sec}^0_U(\mathcal{B}_{C_U}/C_U)_x$ is quasifinite. Since quasifiniteness is a constructible property (see [EGA IV$_3$ 1966, p. 71]) and $g^{-1}(T_0) \cap U$ is dense in $U$ (because $g$ is finite and $T_0$ is dense in $H_0$), this implies that the scheme $\mathrm{Sec}^0_U(\mathcal{B}_{C_U}/C_U)$ is quasifinite over an open subset of $U$. Now replace $U$ by one of its open subschemes so that $\mathrm{Sec}^0_U(\mathcal{B}_{C_U}/C_U)$ becomes quasifinite over $U$. Let $b_{00}$ be an upper bound for the cardinality of the fibres of $\mathrm{Sec}^0_U(\mathcal{B}_{C_U}/C_U) \to U$. Using (a), we conclude that we have

$$\#(n_0 \cdot \mathrm{Tor}(\mathcal{B}_0(\kappa(x)))) \leq b_{00}$$

for all $x \in U(t_0)$. In particular $b_{00}! \cdot n_0 \cdot \mathrm{Tor}(\mathcal{B}_0(\kappa(x)))$ is the trivial group. Thus by the structure of finite subgroups of abelian varieties, we have

$$\#\mathrm{Tor}(\mathcal{B}_0(\kappa(x)))) \leq (b_{00}! \cdot n_0)^{2\dim(\mathcal{B}_{C_U}/C_U)},$$

and we choose $b_0 := (b_{00}! \cdot n_0)^{2\dim(\mathcal{B}_{C_U}/C_U)}$. Finally we let as before $V$ be the open set $H_0 \backslash g(H \backslash U)$. By construction, we have $g^{-1}(V) \subseteq U$. Thus every point of $V(t_0)$ lifts to a point of $U(t_0)$ (since $g$ is finite) and we see that $V$ has the required properties. □

## Appendix B: Ampleness of the Hodge bundle and inseparable points

The terminology of this appendix is independent of the terminology of the rest of the article and appendices. In this appendix, we shall prove a mild extension of the main result of [Rössler 2015].

Let $k$ be a perfect field and let $S$ be a geometrically connected, smooth and proper curve over $k$. Let $K := \kappa(S)$ be its function field. Suppose from now on that $k$ has characteristic $p > 0$.

Let $\pi : \mathcal{A} \to S$ be a smooth commutative group scheme and let $A := \mathcal{A}_K$ be the generic fibre of $\mathcal{A}$. Let $\epsilon_{\mathcal{A}/S} : S \to \mathcal{A}$ be the zero-section and let $\omega := \epsilon^*_{\mathcal{A}/S}(\Omega^1_{\mathcal{A}/S})$ be the Hodge bundle of $\mathcal{A}$ over $S$.

**Theorem B.1.** *Suppose that $\mathcal{A}/S$ is semiabelian, that $A$ is an abelian variety and that $\bar{\mu}_{\min}(\omega) > 0$. Then there exists $\ell_0 \in \mathbb{N}$ such the natural injection $A(K^{p^{-\ell_0}}) \hookrightarrow A(K^{\mathrm{perf}})$ is surjective (and hence a bijection).*

N.B. In [Rössler 2015, Theorem 1.1], Theorem B.1 was proven under the assumption that $A$ is principally polarised and that $k$ is algebraically closed. In can be shown that the condition $\bar{\mu}_{\min}(\omega) > 0$ is equivalent to the requirement that $\omega$ is an ample bundle (see [Rössler 2015, Introduction] for detailed references).

*Proof.* Notice first that in our proof of Theorem B.1, we may replace $K$ by a finite extension field $K'$ without restriction of generality. We may thus suppose that $A$ is endowed with an $m$-level structure for some $m \geqslant 3$ with $(m, p) = 1$.

If $Z \to W$ is a $W$-scheme and $W$ is a scheme of characteristic $p$, then for any $n \geqslant 0$ we shall write $Z^{[n]} \to W$ for the $W$-scheme given by the composition of arrows

$$Z \to W \xrightarrow{F_W^n} W.$$

Now fix $n \geqslant 1$ and suppose that $A(K^{p^{-n}}) \backslash A(K^{p^{-n+1}}) \neq \varnothing$.

Fix $P \in A^{(p^n)}(K) \backslash A^{(p^{n-1})}(K) = A(K^{p^{-n}}) \backslash A(K^{p^{-n+1}})$. The point $P$ corresponds to a commutative diagram of $k$-schemes

$$\begin{array}{ccc} & & A \\ & \overset{P}{\nearrow} & \downarrow \\ \operatorname{Spec} K^{[n]} & \xrightarrow{F_K^n} & \operatorname{Spec} K \end{array}$$

such that the residue field extension $K|\kappa(P(\operatorname{Spec} K^{[n]}))$ is of degree 1 (in other words $P$ is birational onto its image). In particular, the map of $K$-vector spaces $P^*(\Omega^1_{A/k}) \to \Omega^1_{K^{[n]}/k}$ arising from the diagram is nonzero.

Now recall that there is a canonical exact sequence

$$0 \to \pi_K^*(\Omega^1_{K/k}) \to \Omega^1_{A/k} \to \Omega^1_{A/K} \to 0.$$

Furthermore the map $F_K^{n,*}(\Omega^1_{K/k}) \xrightarrow{F_K^{n,*}} \Omega^1_{K^{[n]}/k}$ vanishes. Also, we have a canonical identification $\Omega^1_{A/K} = \pi_K^*(\omega_K)$ (see [Bosch et al. 1990, Chapter 4, Proposition 2]). Thus the natural surjection $P^*(\Omega^1_{A/k}) \to \Omega^1_{K^{[n]}/k}$ gives rise to a nonzero map

$$\phi_n = \phi_{n,P} : F_K^{n,*}(\omega_K) \to \Omega^1_{K^{[n]}/k}.$$

The next crucial lemma examines the poles of the morphism $\phi_n$.

We let $E$ be the reduced closed subset, which is the union of the points $s \in S$, such that the fibre $\mathcal{A}_s$ is not complete.

**Lemma B.2.** *The morphism $\phi_n$ extends to a morphism of vector bundles*

$$F_S^{n,*}(\omega) \to \Omega^1_{S^{[n]}/k}(E).$$

*Proof of Lemma B.2.* First notice that there is a natural identification $\Omega^1_{S^{[n]}/k}(\log E) = \Omega^1_{S^{[n]}/k}(E)$, because there is a sequence of coherent sheaves

$$0 \to \Omega_{S^{[n]}/k} \to \Omega^1_{S^{[n]}/k}(\log E) \to \mathcal{O}_E \to 0,$$

where the morphism onto $\mathcal{O}_E$ is the residue morphism. Here the sheaf $\Omega^1_{S^{[n]}/k}(\log E)$ is the sheaf of differentials on $S^{[n]} \backslash E$ with logarithmic singularities along $E$. See [Illusie 1990, Introduction] for this result and more details on these notions.

We may also suppose without restriction of generality that $A$ is principally polarised. Indeed, consider the following reasoning. By Zarhin's trick, the abelian variety $B := (A \times_K A^\vee)^4$ is principally polarised. Also, $B$ can be endowed with an $m$-level structure compatible with the given $m$-level structure on $A$,

since $A^\vee$ is isogenous to $A$. Let $\mathcal{B} := (\mathcal{A} \times_K \mathcal{A}^\vee)^4$, where (abusing language) we have written $\mathcal{A}^\vee$ for the connected component of the zero-section of the Néron model of $A^\vee$. The group scheme $\mathcal{A}^\vee$ is also semiabelian, since $A^\vee$ is isogenous to $A$ over $K$. The morphism $P \times 0 \times 0 \times \cdots \times 0$ (seven times) gives a point in $B^{(p^n)}(K)$ and there is a commutative diagram

$$\begin{array}{ccc} F_K^{n,*}(\omega_{\mathcal{B},K}) & \xrightarrow{\phi_{n,P\times 0\times\cdots}} & \Omega^1_{K^{[n]}/k} \\ \downarrow & & = \uparrow \\ F_K^{n,*}(\omega_{\mathcal{A},K}) & \xrightarrow{\phi_{n,P}} & \Omega^1_{K^{[n]}/k} \end{array} \tag{12}$$

where the vertical arrow on the left is the pull-back map induced by the closed immersion $\lambda \mapsto \lambda \times 0 \times 0 \times \cdots \times 0$ (seven times). Now since $B$ is principally polarised, we know that if Lemma B.2 holds for principally polarised abelian varieties, the upper row of the diagram (12) extends to a morphism $F_S^{n,*}(\omega_{\mathcal{B}}) \to \Omega^1_{S^{[n]}/k}(E)$ (note that the set of points, where $\mathcal{B}$ is not complete coincides with the set of points, where $\mathcal{A}$ is not complete). Since $F_S^{n,*}(\omega_{\mathcal{A}})$ is a direct summand of $F_S^{n,*}(\omega_{\mathcal{B}})$, we see that Lemma B.2 holds for $A$ if it holds for $B$, thus completing the reduction of Lemma B.2 to the principally polarised case. □

The rest of the proof of Theorem B.1 is identical word for word with the proof of Theorem 1.1 in [Rössler 2015] (from the beginning of the proof of Lemma 2.1). □

## Appendix C: Specialisation of the Mordell–Weil group

The terminology of this appendix is independent of the terminology of the rest of the article and appendices.

In this appendix, we shall prove a geometric analogue of Néron's result on the specialisation of the generic Mordell–Weil group to a fibre in a family of abelian varieties over number fields (see [Lang 1983, Chapter 9, Corollary 6.3]). The following results are reminiscent of some results proven by Hrushovski [1998] in a mixed characteristic context and they are probably already known to many people but we include complete proofs for lack of a reference.

Let $l_0$ be an algebraically closed field. Let $U$ be a smooth and connected quasiprojective variety over $l_0$. Let $\mathcal{B}$ be an abelian scheme over $U$. Suppose given an immersion $\iota : U \hookrightarrow \mathbb{P}^N$ for some $N \geq 0$. Let $K$ be the function field of $U$ and let $B := \mathcal{B}_K$.

**Proposition C.1.** *Suppose that $\mathcal{B}(U)$ is finitely generated. For almost all linear subspaces $L \subseteq \mathbb{P}^N$ of codimension* $\dim(U) - 1$, *the intersection $C := L \cap U$ is smooth*, *connected*, *nonempty*, *the specialisation map*

$$\mathcal{B}(U) \to \mathcal{B}_C(C)$$

*is injective and* $\mathrm{Tr}_{\kappa(C)|l_0}(\mathcal{B}_{\kappa(C)}) = 0$.

Recall that the linear subspaces $L \subseteq \mathbb{P}^N$ of codimension $\dim(U) - 1$ are classified by the Grassmannian $\mathrm{Gr}(\dim(U) - 1, N)$, which is smooth and projective over $l_0$. The words "almost all" stand for "for all the $l_0$-rational points of some dense Zariski open subset of $\mathrm{Gr}(\dim(U) - 1, N)$".

Recall that by a theorem of Weil, the restriction map $\mathcal{B}(U) \to B(K)$ is a bijection. Thus, by the Lang–Néron theorem, the condition that $\mathcal{B}(U) = B(K)$ is finitely generated is equivalent to the condition $\mathrm{Tr}_{K|l_0}(B) = 0$.

For the proof of Proposition C.1, we shall need a few lemmata:

**Lemma C.2.** *Let $N$ be a finite étale group scheme over $U$. Let $t \in H^1_{\mathrm{et}}(U, N)$ and suppose that $t \neq 0$. Then for almost all linear subspaces $L \subseteq \mathbb{P}^N$ of codimension $\dim(U) - 1$, the intersection $C := L \cap U$ is smooth, connected, nonempty and the restriction $t_C \in H^1_{\mathrm{et}}(C, N_C)$ of $t$ to $C$ does not vanish.*

*Proof.* Let $T \to U$ be a torsor under $N$. Note that the torsor $T$ is nontrivial if and only if for all the irreducible components $T'$ of $T$, the (automatically flat and finite) morphism $T' \to U$ has degree $> 1$. The same remark applies to the restriction of $T$ to a smooth and connected closed subscheme of $U$.

Let $(T_i)$ be the set of irreducible components of $T$.

By Bertini's theorem in Jouanolou's presentation [1983, p. 89, Corollary 6.11], for almost all linear subspaces $L \subseteq \mathbb{P}^N$ of codimension $\dim(U) - 1$,

- the intersection $C := L \cap U$ is smooth, connected and nonempty;
- all the $T_{i,C}$ are irreducible.

Let $C$ be in this class. Suppose that $T \to U$ is not trivial. By construction, the irreducible components of $T_C$ are the $T_{i,C}$. Since $T_{i,C} \to C$ is flat and finite of the same degree as $T_i \to U$, we see that the irreducible components of $T_C$ all have degree $> 1$ over $C$. Hence the torsor $T_C$ is not trivial. □

**Lemma C.3.** *Let $N$ be a finite étale group scheme over $U$. Suppose that $N(U) = 0$. Then for almost all linear subspaces $L \subseteq \mathbb{P}^N$ of codimension $\dim(U) - 1$, the intersection $C := L \cap U$ is smooth, connected, nonempty and $N_C(C) = 0$.*

*Proof.* Let $(N_i)$ be the set of irreducible components of $N$, excluding the component of the identity. The condition that $N(U) = 0$ is equivalent to the condition that for all $i$, the morphism $N_i \to U$ has degree $> 1$.

As before, by Bertini's theorem, for almost all linear subspaces $L \subseteq \mathbb{P}^N$ of codimension $\dim(U) - 1$,

- the intersection $C := L \cap U$ is smooth and connected;
- all the $N_{i,C}$ are irreducible.

Let $C$ be in this class. By construction, the irreducible components of $N_C$ outside of the component of the identity are the $N_{i,C}$. Since $N_{i,C} \to C$ is flat and finite of the same degree as $N_i \to U$, we see that the irreducible components of $N_C$ outside of the component of the identity all have degree $> 1$ over $C$. Hence $N_C(C) = 0$. □

**Lemma C.4.** *Let $G \subseteq \mathcal{B}(U)$ be a finite group. For almost all linear subspaces $L \subseteq \mathbb{P}^N$ of codimension $\dim(U) - 1$, the intersection $C := L \cap U$ is smooth and connected, and the reduction map*

$$G \to \mathcal{B}_C(C)$$

*is injective.*

*Proof.* The proof is left to the reader. □

Finally, we need an elementary but very insightful lemma, due to in essence to Néron. The following version is due to Hrushovski [1998, Lemma 1]:

**Lemma C.5** (Néron–Hrushovski). *Let $r : G \to H$ be a map of abelian groups. Let $l$ be a prime number. Suppose that* $\mathrm{Tor}_l(H) = 0$ *and that the induced map $G/lG \to H/lH$ is injective. Then* $\ker r \subseteq \bigcap_{j\geq 0} l^j G$.

*Proof.* Let $g \in \ker r$. Suppose for contradiction that $g \notin \bigcap_{j\geq 0} l^j G$. Let $m \geq 0$ be the smallest natural number such that $g \notin l^m G$. Then there is $g' \in G$ such that $l^{m-1}g' = g$ and thus $r(g') \in \mathrm{Tor}_l(H)$ so that from the assumptions we have $r(g') = 0$. Since the map $G/lG \to H/lH$ is injective, there is $g'' \in G$ such that $lg'' = g'$. Hence $g = l^m g''$, a contradiction. □

*Proof of Proposition C.1.* Let $l$ be a prime number such that $\mathrm{Tor}_l(\mathcal{B}(U)) = 0$ and such that $l$ is not the characteristic of $l_0$. Notice that for any closed subscheme $C$ of $U$, we have an injection $\delta_C : \mathcal{B}(C)/l\mathcal{B}(C) \hookrightarrow H^1_{\mathrm{et}}(C, \ker[l]_{\mathcal{B},C})$ and this injection is functorial for restrictions to smaller closed subschemes $C_1 \hookrightarrow C$. According to Lemmata C.2, C.3 and C.4, for almost all linear subspaces $L \subseteq \mathbb{P}^N$ of codimension $\dim(U) - 1$,

- the intersection $C := L \cap U$ is smooth and connected;
- the restriction map $H^1(U, \ker[l]_{\mathcal{B}}) \to H^1(C, \ker[l]_{\mathcal{B},C})$ is injective on the image of $\delta_U$;
- $(\ker[l]_{\mathcal{B},C})(C) = 0$;
- the restriction map $\mathrm{Tor}(\mathcal{B}(U)) \to \mathcal{B}(C)$ is injective.

Let $C$ be in this class. By construction, the map $\mathcal{B}(U)/l\mathcal{B}(U) \to \mathcal{B}(C)/l\mathcal{B}(C)$ is injective and $\mathrm{Tor}_l(\mathcal{B}(C)) = 0$. Let $F$ be a free subgroup of $\mathcal{B}(U)$, which is a direct summand of $\mathrm{Tor}(\mathcal{B}(U))$. We have $F \cap \left(\bigcap_{j\geq 0} l^j \mathcal{B}(U)\right) = 0$ since $\mathcal{B}(U)$ is finitely generated and $F$ is free. Applying Lemma C.5 to $G = \mathcal{B}(U)$ and $H = \mathcal{B}(C)$, we see that the restriction map $F \to \mathcal{B}(C)$ is injective. Since the restriction map $\mathrm{Tor}(\mathcal{B}(U)) \to \mathcal{B}(C)$ is also injective, we thus see that the restriction map $\mathcal{B}(U) \to \mathcal{B}(C)$ is injective. Finally, we have $\mathrm{Tr}_{\kappa(C)|l_0}(\mathcal{B}_{\kappa(C)}) = 0$, for otherwise, we would have $\mathrm{Tor}_l(\mathcal{B}(C)) \neq 0$. □

## Acknowledgments

My warm thanks to the referee for his/her careful reading and for many suggestions. The article would be much less clear without his/her help and encouragement. I would like to thank J.-B. Bost for his feedback, especially for pointing out the article [Catanese and Dettweiler 2016], for suggesting Remark 2.4 and for providing [Bost 2004, Lemma 2.9], whose positive characteristic analogue is technically at the root of the present text. Minhyong Kim's article [1997] also played a fundamental role in the genesis of the present text; the construction described there pointed me in (what I hope is) the right direction when I started studying purely inseparable points on abelian varieties. I had many interesting discussions with him about his article. I am very grateful to J.-F. Voloch for many exchanges on the material of this article and for his remarks on the text and to P. Ziegler for many discussions on and around the "full" Mordell–Lang conjecture. Many thanks also to T. Scanlon for his interest and for interesting discussions

around the group $A(K^{\mathrm{perf}})$. Last but not least, many thanks to Hélène Esnault and her student Marco d'Addezio for their interest and for many enlightening discussions around Theorem 1.4. I also benefitted from A.-J. de Jong's and F. Oort's vast knowledge; they both very kindly took the time to answer some rather speculative messages.

## References

[Abbes 2000] A. Abbes, "Réduction semi-stable des courbes d'après Artin, Deligne, Grothendieck, Mumford, Saito, Winters, . . .", pp. 59–110 in *Courbes semi-stables et groupe fondamental en géométrie algébrique* (Luminy, 1998), edited by J.-B. Bost et al., Progr. Math. **187**, Birkhäuser, Basel, 2000. MR Zbl

[Abramovich and Voloch 1992] D. Abramovich and J. F. Voloch, "Toward a proof of the Mordell–Lang conjecture in characteristic $p$", *Int. Math. Res. Not.* 5 (1992), 103–115. MR Zbl

[Artin and Milne 1976] M. Artin and J. S. Milne, "Duality in the flat cohomology of curves", *Invent. Math.* **35** (1976), 111–129. MR Zbl

[Barton 1971] C. M. Barton, "Tensor products of ample vector bundles in characteristic $p$", *Amer. J. Math.* **93** (1971), 429–438. MR Zbl

[Bosch et al. 1990] S. Bosch, W. Lütkebohmert, and M. Raynaud, *Néron models*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) **21**, Springer, 1990. MR Zbl

[Bost 2004] J.-B. Bost, "Germs of analytic varieties in algebraic varieties: canonical metrics and arithmetic algebraization theorems", pp. 371–418 in *Geometric aspects of Dwork theory*, vol. I, edited by A. Adolphson et al., Walter de Gruyter, Berlin, 2004. MR Zbl

[Brenner et al. 2008] H. Brenner, J. Herzog, and O. Villamayor, *Three lectures on commutative algebra*, edited by G. Colomé-Nin et al., University Lecture Series **42**, Amer. Math. Soc., Providence, RI, 2008. MR Zbl

[Catanese and Dettweiler 2016] F. Catanese and M. Dettweiler, "Vector bundles on curves coming from variation of Hodge structures", *Internat. J. Math.* **27**:7 (2016), art. id. 1640001, 25 pp. MR Zbl

[Conrad 2006] B. Conrad, "Chow's $K/k$-image and $K/k$-trace, and the Lang–Néron theorem", *Enseign. Math.* (2) **52**:1-2 (2006), 37–108. MR Zbl

[EGA IV$_2$ 1965] A. Grothendieck, "Eléments de géométrie algébrique, IV: Étude locale des schémas et des morphismes de schémas, II", *Inst. Hautes Études Sci. Publ. Math.* **24** (1965), 5–231. MR Zbl

[EGA IV$_3$ 1966] A. Grothendieck, "Eléments de géométrie algébrique, IV: Étude locale des schémas et des morphismes de schémas, III", *Inst. Hautes Études Sci. Publ. Math.* **28** (1966), 5–255. MR Zbl

[EGA IV$_4$ 1967] A. Grothendieck, "Eléments de géométrie algébrique, IV: Étude locale des schémas et des morphismes de schémas, IV", *Inst. Hautes Études Sci. Publ. Math.* **32** (1967), 5–361. MR Zbl

[Ekedahl 1988] T. Ekedahl, "Canonical models of surfaces of general type in positive characteristic", *Inst. Hautes Études Sci. Publ. Math.* **67** (1988), 97–144. MR Zbl

[Esnault and Langer 2013] H. Esnault and A. Langer, "On a positive equicharacteristic variant of the $p$-curvature conjecture", *Doc. Math.* **18** (2013), 23–50. MR Zbl

[Faltings and Chai 1990] G. Faltings and C.-L. Chai, *Degeneration of abelian varieties*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) **22**, Springer, 1990. MR Zbl

[Fantechi et al. 2005] B. Fantechi, L. Göttsche, L. Illusie, S. L. Kleiman, N. Nitsure, and A. Vistoli, *Fundamental algebraic geometry: Grothendieck's FGA explained*, Mathematical Surveys and Monographs **123**, Amer. Math. Soc., Providence, RI, 2005. MR Zbl

[Ghioca and Moosa 2006] D. Ghioca and R. Moosa, "Division points on subvarieties of isotrivial semi-abelian varieties", *Int. Math. Res. Not.* **2006** (2006), art. id. 65437, 23 pp. MR Zbl

[Grothendieck 1966] A. Grothendieck, "Techniques de construction et théorèmes d'existence en géométrie algébrique IV: Les schémas de Hilbert", exposé 221 in *Séminaire Bourbaki,* 1960/1961, W. A. Benjamin, Amsterdam, 1966. Reprinted as pp. 249–276 in *Séminaire Bourbaki* **6**, Soc. Math. France, Paris, 1995. MR

[Grothendieck 1974] A. Grothendieck, *Groupes de Barsotti–Tate et cristaux de Dieudonné* (Montréal, 1970), Les Presses de l'Université de Montréal, 1974. MR Zbl

[Hartshorne 1977] R. Hartshorne, *Algebraic geometry*, Graduate Texts in Mathematics **52**, Springer, 1977. MR Zbl

[Hrushovski 1998] E. Hrushovski, "Proof of Manin's theorem by reduction to positive characteristic", pp. 197–205 in *Model theory and algebraic geometry*, edited by E. Bouscaren, Lecture Notes in Math. **1696**, Springer, 1998. MR Zbl

[Illusie 1990] L. Illusie, "Réduction semi-stable et décomposition de complexes de de Rham à coefficients", *Duke Math. J.* **60**:1 (1990), 139–185. MR Zbl

[Jouanolou 1983] J.-P. Jouanolou, *Théorèmes de Bertini et applications*, Progress in Mathematics **42**, Birkhäuser, Boston, 1983. MR Zbl

[Katz 1981] N. Katz, "Serre-Tate local moduli", pp. 138–202 in *Algebraic surfaces* (Orsay, 1976–78), edited by J. Giraud et al., Lecture Notes in Math. **868**, Springer, 1981. MR Zbl

[Kim 1997] M. Kim, "Purely inseparable points on curves of higher genus", *Math. Res. Lett.* **4**:5 (1997), 663–666. MR Zbl

[Künnemann 1998] K. Künnemann, "Projective regular models for abelian varieties, semistable reduction, and the height pairing", *Duke Math. J.* **95**:1 (1998), 161–212. MR Zbl

[Lang 1983] S. Lang, *Fundamentals of Diophantine geometry*, Springer, 1983. MR Zbl

[Langer 2004] A. Langer, "Semistable sheaves in positive characteristic", *Ann. of Math.* (2) **159**:1 (2004), 251–276. MR Zbl

[Langer 2015] A. Langer, "Generic positivity and foliations in positive characteristic", *Adv. Math.* **277** (2015), 1–23. MR Zbl

[Liu 2002] Q. Liu, *Algebraic geometry and arithmetic curves*, Oxford Graduate Texts in Mathematics **6**, Oxford University Press, 2002. MR Zbl

[Milne 1980] J. S. Milne, *Étale cohomology*, Princeton Mathematical Series **33**, Princeton University Press, 1980. MR Zbl

[Milne 1986] J. S. Milne, "Abelian varieties", pp. 103–150 in *Arithmetic geometry* (Storrs, CT, 1984), edited by G. Cornell and J. H. Silverman, Springer, 1986. MR Zbl

[Milne 2006] J. S. Milne, *Arithmetic duality theorems*, 2nd ed., BookSurge, Charleston, SC, 2006. MR Zbl

[Milne 2017] J. S. Milne, *Algebraic groups: the theory of group schemes of finite type over a field*, Cambridge Studies in Advanced Mathematics **170**, Cambridge University Press, 2017. MR Zbl

[Moret-Bailly 1985] L. Moret-Bailly, *Pinceaux de variétés abéliennes*, Astérisque **129**, Société Mathématique de France, Paris, 1985. MR Zbl

[Mumford 1970] D. Mumford, *Abelian varieties*, Tata Institute of Fundamental Research Studies in Mathematics **5**, Oxford University Press, London, 1970. MR Zbl

[Mumford et al. 1994] D. Mumford, J. Fogarty, and F. Kirwan, *Geometric invariant theory*, 3rd ed., Ergebnisse der Mathematik und ihrer Grenzgebiete (2) **34**, Springer, 1994. MR Zbl

[Nitsure 2005] N. Nitsure, "Construction of Hilbert and Quot schemes", pp. 105–137 in *Fundamental algebraic geometry: Grothendieck's FGA explained*, Math. Surveys Monogr. **123**, Amer. Math. Soc., Providence, RI, 2005. MR

[Poonen and Voloch 2010] B. Poonen and J. F. Voloch, "The Brauer–Manin obstruction for subvarieties of abelian varieties over function fields", *Ann. of Math.* (2) **171**:1 (2010), 511–532. MR Zbl

[Raynaud 1970] M. Raynaud, *Faisceaux amples sur les schémas en groupes et les espaces homogènes*, Lecture Notes in Mathematics **119**, Springer, 1970. MR Zbl

[Rössler 2013] D. Rössler, "Infinitely $p$-divisible points on abelian varieties defined over function fields of characteristic $p > 0$", *Notre Dame J. Form. Log.* **54**:3-4 (2013), 579–589. MR Zbl

[Rössler 2015] D. Rössler, "On the group of purely inseparable points of an abelian variety defined over a function field of positive characteristic", *Comment. Math. Helv.* **90**:1 (2015), 23–32. MR Zbl

[Rössler 2019a] D. Rössler, "Le groupe de Selmer des isogénies de hauteur un", preprint, 2019. arXiv

[Rössler 2019b] D. Rössler, "Perfect points on abelian varieties in positive characteristic", video, 2019, available at https://www.youtube.com/watch?v=TKLEIpNsyDI.

[Scanlon 2005] T. Scanlon, "A positive characteristic Manin–Mumford theorem", *Compos. Math.* **141**:6 (2005), 1351–1364. MR Zbl

[Serre and Tate 1968] J.-P. Serre and J. Tate, "Good reduction of abelian varieties", *Ann. of Math.* (2) **88** (1968), 492–517. MR Zbl

[SGA $3_{\mathrm{I}}$ 2011] M. Demazure and A. Grothendieck, *Schémas en groupes, Tome I: Propriétés générales des schémas en groupes, Exposés I–VII* (Séminaire de Géométrie Algébrique du Bois Marie 1962–1964), revised ed., edited by P. Gille and P. Polo, Documents Mathématiques (Paris) **7**, Société Mathématique de France, Paris, 2011. MR Zbl

[SGA $3_{\mathrm{II}}$ 1970] M. Demazure and A. Grothendieck, *Schémas en groupes, Tome II: Groupes de type multiplicatif, et structure des schémas en groupes généraux, Exposés VIII–XVIII* (Séminaire de Géométrie Algébrique du Bois Marie 1962–1964), Lecture Notes in Math. **152**, Springer, 1970. MR Zbl

[Shepherd-Barron 1992] N. I. Shepherd-Barron, "Miyaoka's theorems on the generic seminegativity of $T_X$ and on the Kodaira dimension of minimal regular threefolds", pp. 103–114 in *Flips and abundance for algebraic threefolds*, Astérisque **211**, Société Mathématique de France, Paris, 1992. Zbl

[Szamuely 2010] T. Szamuely, "Corps de classes des schémas arithmétiques", exposé 1006, pp. 257–286 in *Séminaire Bourbaki,* 2008/2009, Astérisque **332**, Société Mathématique de France, Paris, 2010. MR Zbl

[Tate 1997] J. Tate, "Finite flat group schemes", pp. 121–154 in *Modular forms and Fermat's last theorem* (Boston, MA, 1995), edited by G. Cornell et al., Springer, 1997. MR Zbl

[Tate and Oort 1970] J. Tate and F. Oort, "Group schemes of prime order", *Ann. Sci. École Norm. Sup.* (4) **3** (1970), 1–21. MR Zbl

[Voloch 1995] J. F. Voloch, "Diophantine approximation on abelian varieties in characteristic $p$", *Amer. J. Math.* **117**:4 (1995), 1089–1095. MR Zbl

[Yuan 2018] X. Yuan, "Positivity of Hodge bundles of abelian varieties over some function fields", preprint, 2018. arXiv

[Zarkhin and Parshin 1989] Y. G. Zarkhin and A. N. Parshin, "Finiteness problems in Diophantine geometry", pp. 35–102 in *American Mathematical Society translations*, edited by B. Silver, American Mathematical Society Translations (2) **143**, Amer. Math. Soc., Providence, RI, 1989. Zbl

Communicated by Hélène Esnault
Received 2018-09-20 Revised 2019-11-19 Accepted 2019-12-17

damian.rossler@maths.ox.ac.uk *Mathematical Institute, University of Oxford, Oxford, United Kingdom*

ALGEBRA AND NUMBER THEORY 14:5 (2020)
https://doi.org/10.2140/ant.2020.14.1175

# Mixed Tate motives and the unit equation II

Ishai Dan-Cohen

Over the past fifteen years or so, Minhyong Kim has developed a framework for making effective use of the fundamental group to bound (or even compute) integral points on hyperbolic curves. This is the third installment in a series whose goal is to realize the potential effectivity of Kim's approach in the case of the thrice punctured line. As envisioned by Dan-Coehn and Wewers (2016), we construct an algorithm whose output upon halting is provably the set of integral points, and whose halting would follow from certain natural conjectures. Our results go a long way towards achieving our goals over the rationals, while broaching the topic of higher number fields.



## 1. Introduction

**1.1.** This is the third installment in a series [Dan-Cohen and Wewers 2015; 2016][1] devoted to what may reasonably be described as *explicit motivic Chabauty–Kim theory*. "Chabauty–Kim theory" refers to a framework developed by Minhyong Kim for making effective use of the fundamental group to bound, or conjecturally compute, integral solutions to hyperbolic equations. "Motivic" refers to the fact that while Kim's construction, in its original formulation, is $p$-adic étale, our methods are motivic. As things currently stand, this limits us to working in the mixed Tate, or Artin–Tate settings, that is, essentially to the projective line with (possibly interesting) punctures. So the adjective "motivic" implies a fairly

This work was supported by Priority Program 1489 of the Deutsche Forschungsgemeinschaft: *Experimental and algorithmic methods in algebra, geometry, and number theory*.
*MSC2010:* primary 11G55; secondary 11D45, 14F30, 14F35, 14F42, 14G05.
*Keywords:* mixed Tate motives, unipotent fundamental group, p-adic periods, polylogarithms, unit equation, integral points.

[1] *Explicit Chabauty–Kim theory for the thrice punctured line in depth two = Mixed Tate motives and the unit equation 0.*

specific context. While this context may seem narrow from a geometric point of view, it is quite broad from an arithmetic point of view, leading and relating to various interesting questions and conjectures.

"Explicit" refers to the fact that here our emphasis is on algorithms. *Explicit* Chabauty–Kim theory, as I see it, is somewhat orthogonal to Chabauty–Kim theory proper. If *Chabauty–Kim theory* is about attempting to prove Kim's conjecture [Balakrishnan et al. 2018], or at least about formulating and studying a range of related conjectures, *Explicit* Chabauty–Kim theory is about making the theory *explicit*. In particular, in the explicit theory, we allow ourselves to assume conjectures left and right, so long as those affect the halting, and not the construction, of the hoped-for algorithms.

In this installment, we continue our study of $\mathbb{P}^1 \setminus \{0, 1, \infty\}$. We obtain an algorithm for computing the *polylogarithmic Chabauty–Kim loci* (see below) over number fields which obey a certain technical condition. In turn, this technical condition is known for the rationals and follows for general number fields from a conjecture due to Jannsen. We also obtain an algorithmic solution to the unit equation over totally real fields obeying the same condition. Specializing to the case of the rationals, we obtain the algorithm envisioned by Dan-Cohen and Wewers [2016].

**1.2.** We now state our main application in more detail. Let $X = \mathbb{P}^1 \setminus \{0, 1, \infty\}$. Below, we construct an algorithm which takes as input an open integer scheme $Z$ (by which we mean an open subscheme of $\operatorname{Spec} \mathcal{O}_K$ for $K$ a number field), and outputs a subset of $X(Z)$.

**Theorem 1.2.1.** *Let $Z$ be a totally real open integer scheme. If our algorithm halts for the input $Z$, then its output is equal to the set $X(Z)$ of integral points of $X$ over $Z$.*

We also state four conjectures: *Zagier's conjecture*, *Goncharov exhaustion* (*with weak control over ramification*), the *p-adic period conjecture*, and *convergence of Chabauty–Kim loci for the polylogarithmic quotient*. Finally, we state our technical condition, which we call *Hasse principle for finite cohomology*. We say that $K$ obeys *Kim vs. Hasse* if convergence occurs *before* the Hasse principle fails (see below for details). The following proposition motivates the theorem above.

**Proposition 1.2.2.** *Let $Z$ be a totally real open integer scheme with fraction field $K$*:

(1) *Assume Zagier's conjecture*, *Goncharov exhaustion*, *the p-adic period conjecture*, *and convergence of Chabauty–Kim loci for the polylogarithmic quotient hold for $Z$. Assume $K$ obeys Kim vs. Hasse. Then our algorithm halts for $Z$.*

(2) *Suppose $Z$ is contained in* $\operatorname{Spec} \mathbb{Z}$. *Assume Goncharov exhaustion*, *the p-adic period conjecture*, *and convergence of Chabauty–Kim loci for the polylogarithmic quotient hold for $Z$. Then our algorithm halts for $Z$.*

We refer to Theorem 1.2.1 and Proposition 1.2.2, taken together, as the "equation-solving theorem"; see Theorem 7.2.1 for a more precise statement.

Practical (and unconditional) methods for solving the $S$-unit equation predate this work, and can be found, for instance, in de Weger [1989] who uses the theory of logarithmic forms of Baker and Wüsthotz [2007] (see also [Evertse and Győry 2015] for a general discussion). A more recent approach, due to von

Känel and Matschke [2016] is based on the Shimura–Taniyama conjecture. Our primary purpose here is not to compete with these other methods, but rather, to develop Kim's theory in a special case, to explore its interaction with the theory of mixed Tate motives and motivic iterated integrals, and, in subsequent works, to provide new numerical evidence for Kim's conjecture. Also, while our focus here is *explicit*, it is not yet *practical*; see segment 1.15.1 below.

**1.3.** We now give a brief indication of our main result, from which the equation-solving theorem follows as a corollary; precise statements, as well as more background, appear in Section 2 below. For this purpose, fix a prime $\mathfrak{p} \in Z$ which we assume to be totally split and recall that there is a commutative diagram like so:

$$\begin{array}{ccc} X(Z) & \longrightarrow & X(Z_{\mathfrak{p}}) \\ \downarrow{\scriptstyle \kappa} & & \downarrow{\scriptstyle \kappa_{\mathfrak{p}}} \\ \mathbb{Q}_p \otimes H^1(U^{\mathrm{PL}}_{\geq -n}) & \xrightarrow[\mathfrak{R}_{\mathfrak{p}}]{} & H^1(U^{\mathrm{PL},F\phi}_{\geq -n}) \end{array}$$

Here $Z_{\mathfrak{p}}$ denotes the complete local scheme of $Z$ at $\mathfrak{p}$, isomorphic to $\operatorname{Spec}\mathbb{Z}_p$, and $U^{\mathrm{PL}}_{\geq -n}$ denotes the level-$n$ quotient of the polylogarithmic quotient of the unipotent fundamental group of $X$ at the tangential base point $\vec{1}_0$, a certain quotient of a unipotent, motivic version of the fundamental group. The cohomology variety $H^1(U^{\mathrm{PL}}_{\geq -n})$ appearing below left is a certain $\mathbb{Q}$-variety parametrizing torsors for $U^{\mathrm{PL}}_{\geq -n}$. The vertical map $\kappa$ sends an integral point $x$ to the torsor of homotopy classes of paths

$$1_0 \to x.$$

The cohomology variety appearing in the lower-right is a certain $p$-adic variant of the one to its left, based, as the notation suggests, on the theory of filtered $\phi$ modules. In terms of this diagram, we define

$$X(Z_{\mathfrak{p}})_n := \kappa_{\mathfrak{p}}^{-1}(\operatorname{Im}\mathfrak{R}_{\mathfrak{p}}).$$

We construct an algorithm for computing the locus $X(Z_{\mathfrak{p}})_n$ to given $p$-adic precision.

**Theorem 1.3.1** (see Theorem 2.4.1 below). *Let $Z$ be a totally real open integer scheme, $\mathfrak{p}$ a totally split prime, $n$ a natural number, and $\epsilon > 0$. If our algorithm halts for these inputs, then the functions $\tilde{F}_i^{\mathfrak{p}}$ which the algorithm returns as output take values less than $\epsilon$ on $X(Z_{\mathfrak{p}})_n$.*

**1.4.** The main problem of explicit Chabauty–Kim theory is to render the map $\mathfrak{R}_{\mathfrak{p}}$ computationally accessible; in the case at hand, we proceed as follows. Let $U(Z)$ denote the unipotent part of the fundamental group of the category of mixed Tate motives over $Z$. If $Z^o \subset Z$ is an open subscheme, there's an associated surjection

$$U(Z^o) \twoheadrightarrow U(Z).$$

As part of the algorithm, we search for a $Z^o$ such that $U(Z^o)$ admits a *nice* set of coordinates. More will be said about the role played by $Z^o$ below; for now, let us fix $Z^o$ arbitrarily. A theory of $p$-adic iterated

integration due to Coleman and Besser gives rise to a point

$$I_{\mathrm{BC}} : \operatorname{Spec} \mathbb{Q}_p \to U(Z^o).$$

Our construction revolves around the following diagram:

$$\begin{array}{ccc}
\operatorname{Spec} \mathbb{Q}_p \times H^1(U^{\mathrm{PL}}_{\geq -n}) & \xrightarrow{\quad \mathfrak{R}_{\mathfrak{p}} \quad} & H^1(U^{\mathrm{PL}, F\phi}_{\geq -n}) \\
\| & & \| \\
\operatorname{Spec} \mathbb{Q}_p \times Z^1(U(Z), U^{\mathrm{PL}}_{\geq -n})^{\mathbb{G}_m} & \xrightarrow{\mathfrak{ev}_{I_{\mathrm{BC}}}} & \operatorname{Spec} \mathbb{Q}_p \times U^{\mathrm{PL}}_{\geq -n} \\
{\scriptstyle I_{\mathrm{BC}} \times Id}\downarrow & & \downarrow{\scriptstyle I_{\mathrm{BC}} \times Id} \\
U(Z^o) \times Z^1(U(Z), U^{\mathrm{PL}}_{\geq -n})^{\mathbb{G}_m} & \xrightarrow[\mathfrak{ev}_{\mathrm{Everywhere}}]{} & U(Z^o) \times U^{\mathrm{PL}}_{\geq -n}
\end{array}$$

Here $Z^1(U(Z), U^{\mathrm{PL}}_{\geq -n})^{\mathbb{G}_m}$ denotes a certain space of $\mathbb{G}_m$-equivariant cocycles

$$U(Z) \to U^{\mathrm{PL}}_{\geq -n},$$

and the maps $\mathfrak{ev}_{I_{\mathrm{BC}}}$, $\mathfrak{ev}_{\mathrm{Everywhere}}$ are evaluation maps. Instead of attempting to compute the scheme-theoretic image $\operatorname{Im} \mathfrak{R}_{\mathfrak{p}}$ directly, we compute the pullback

$$(I_{\mathrm{BC}} \times Id)^{-1}(\operatorname{Im} \mathfrak{ev}_{\mathrm{Everywhere}}). \tag{$*$}$$

**1.5.** The group $U(Z^o)$ possesses certain special functions, known as *motivic iterated integrals*, whose pullbacks along $I_{\mathrm{BC}}$ are $p$-adic iterated integrals. The latter may be computed to arbitrary precision thanks to the algorithm of Dan-Cohen and Chatzistamatiou [2014], which we review in Section 6 below. In order to compute $\operatorname{Im} \mathfrak{ev}_{\mathrm{Everywhere}}$ as well as its pullback $(*)$ algorithmically, we need coordinates on $U(Z^o)$. Moreover, as the algorithm proceeds, we need to impose different, in fact contradictory, conditions on our coordinate system: to compute the pullback along $I_{\mathrm{BC}}$, we need our coordinate functions to be given explicitly in terms of motivic iterated integrals. To compute the image

$$\operatorname{Im} \mathfrak{ev}_{\mathrm{Everywhere}}$$

however, we need coordinates compatible with the product on $U(Z^o)$. In the construction that follows, we attempt to bridge this gap; we fail in many ways, but are nevertheless able to make the error incurred arbitrarily small.

This work does not have significant logical dependence on its predecessors [Dan-Cohen and Wewers 2015; 2016]. The reason for this, in part, is that I found it preferable to modify portions of the work done in [Dan-Cohen and Wewers 2016] in preparation for the construction of the algorithm. For instance, our use of the map "$\mathfrak{ev}_{\mathrm{Everywhere}}$" here (along with our acceptance of the $p$-adic period conjecture as yet another condition for halting) allows us to carry out the geometric part of the computation in a single step over the rationals and in a manner entirely divorced from arithmetic. In fact, the relationship between the

present work and the latter is rather reversed: [loc. cit.] may be seen as working out a particular example of the algorithm constructed here.

**1.6.** After making precise the conjectures and theorems indicated above in Section 2, we begin in segements 3.1 and 3.2 by studying formal properties of coordinate systems on $U(Z^o)$ which promise to shrink the apparent gap between computable properties of motivic iterated integrals and the desired compatibility with product. The result, which is summarized in Propositions 3.2.2 and 3.2.3, consists of conditions on a basis $\mathcal{A}$ for the Hopf algebra $A(Z^o)$ of functions on $U(Z^o)$, given as a disjoint union of three subsets

$$\mathcal{A} = \mathcal{E} \cup \mathcal{P} \cup \mathcal{D},$$

under which

$$\mathcal{E} \cup \mathcal{P}$$

forms an algebra basis of the polynomial algebra $A(Z^o)$, and the set $\mathcal{E}^\vee$ of dual elements forms a set of free generators for the Lie algebra

$$\mathfrak{n}(Z^o) := \operatorname{Lie} U(Z^o).$$

Our terminology gives a rough idea of the roles played by these subsets: $\mathcal{E}$ consists of *extensions*, $\mathcal{P}$ of *primitive nonextensions*, and $\mathcal{D}$ of *decomposables*.

**1.7.** Let $U(Z_\mathfrak{p})$ (or $U(\mathcal{O}_\mathfrak{p})$) denote the unipotent part of the fundamental group of the Tannakian category of mixed Tate filtered $\phi$ modules of Chatzistamatiou and Ünver [2013].[2] Let $A(Z_\mathfrak{p})$ denote the *mixed Tate filtered $\phi$ Hopf algebra*, that is, the Hopf algebra of functions on $U(Z_\mathfrak{p})$. Unlike the motivic Galois group $U(Z^\circ)$, the filtered $\phi$ Galois group $U(Z_\mathfrak{p})$ possesses a canonical set of free generators which give rise, dually, to a set of *standard* basis elements in $A(Z_\mathfrak{p})$. There is a *realization map*

$$\mathrm{Re}_p : A(Z^o) \to \prod_{\mathfrak{p} \mid p} A(Z_\mathfrak{p}).$$

Given a motivic iterated integral in $A(Z^o)$, we may wish to expand its realization in the standard basis. In segment 3.7 we upgrade the algorithm of [Dan-Cohen and Chatzistamatiou 2014] to an algorithm which computes $p$-adic approximations of this expansion; we refer to this algorithm as the *realization algorithm*. Examples of this algorithm are worked out in [Dan-Cohen and Wewers 2016, Section 7.5] for $Z = \operatorname{Spec} \mathbb{Z}\left[\frac{1}{2}\right]$:[3]

(1) In segment 7.5.1 we compute, in the notation of that paper, the $p$-adic number $(\log^{F\phi} 2)(v_{-1})$.

(2) In segment 7.5.2 we compute, for instance, $(\log^{F\phi} 2)^2(v_{-2})$.

[2]The same symbols might be used to denote the unipotent part of the fundamental group of the category of mixed Tate motives over $Z_p$ if such a category exists; but this hypothetical group will not intervene in this paper.

[3]Actually, when $Z$ is an open subscheme of $\operatorname{Spec} \mathbb{Z}$ (and the higher motivic extension groups are hence of dimension $\leq 1$), this algorithm may be replaced by a single direct period-computation; see [Corwin and Dan-Cohen 2018a; 2018b].

(3) In segment 7.5.3 we compute $\mathrm{Li}_3^{F\phi}(b)(w)$ for

$$w \in \{v_{-1}^3,\, v_{-1}v_{-2},\, v_{-2}v_{-1},\, v_{-3}\}.$$

**1.8.** Next comes our "basis algorithm" and our "change of basis algorithm". In comparison with the sketch of our algorithm in [Dan-Cohen and Wewers 2016, Section 5.7], the basis algorithm corresponds to step 1 (segment 5.7.1) while the change of basis algorithm is a part of step 2 (segment 5.7.2).[4] This material was inspired by Brown [2012].

In segment 3.8 below we attempt to construct a basis $\mathcal{A}$ of iterated integrals for $A(Z^o)$ (varying $Z^o \subset Z$ as we search) which fulfills the conditions of Proposition 3.2.2. The result is our basis algorithm. Examples can be found in [Dan-Cohen and Wewers 2016, Section 7.5] when we find a basis of $A(Z^o)_n$ for $n = 1, 2, 3, 4$ consisting of unipotent motivic polylogarithms. For instance, in Proposition 7.5.4.1 of [loc. cit.], we find that a basis for $A\big(\mathbb{Z}\big[\frac{1}{2}\big]\big)_4$ is given by

$$\mathcal{B} = \big\{(\log^U 2)^4,\, (\log^U 2)\zeta^U(3),\, \mathrm{Li}_4^U\big(\tfrac{1}{2}\big)\big\}.$$

When constructing the basis algorithm, one problem we face is that we are unable to verify algorithmically if a given iterated integral in $A(Z^o)_r$ belongs to the subspace

$$E(Z^o)_r := \mathrm{Ext}^1_{\boldsymbol{MT}(Z)}(\mathbb{Q}(0), \mathbb{Q}(r)) \subset A(Z^o)_r$$

of extensions. Using the realization algorithm, however, we can bound the distance between a given iterated integral and the extension space, and so bound the error thus incurred. We are thus forced to work with two potentially distinct bases. One basis, denoted by $\tilde{\mathcal{A}}$, is given concretely and explicitly by motivic iterated integrals, but is imperfect in that its set

$$\tilde{\mathcal{E}} \subset \tilde{\mathcal{A}}$$

of alleged extensions may actually fail to be extensions. By projecting $\tilde{\mathcal{E}}$ onto the space of extensions we obtain a second basis, $\mathcal{A}$, which is perfect in its fulfillment of the conditions of Proposition 3.2.2 on the one hand, but is merely *abstract* on the other hand, as its definition is not constructive.

**1.9.** Let us briefly visit segment 3.8.11, where the construction becomes somewhat intricate. The construction is recursive in $n \geq 2$. As soon as we have a basis $\tilde{\mathcal{A}}_{\leq n}$ of motivic iterated integrals in half-weights $\leq n$, we want to also be able to expand an arbitrary iterated integral in half-weight $n$ in the given basis, or, more generally, to compute the inner product of two arbitrary iterated integrals in half-weight $n$ (for the standard inner product $\langle v_i, v_j\rangle = \delta_{i,j}$ induced by this basis). We don't hope to be able to do this precisely; instead we aim for an $\epsilon$-approximation

$$\langle J, I\rangle_\epsilon.$$

[4] Indeed, segment 5.7.2 of [loc. cit.] is a mixture of our *change of basis* algorithm with our "cocycle-image-evaluation algorithm".

Segment 3.8.8 of our algorithm is key in setting the stage for this computation. Under our "Hasse principle", the realization map is injective near the extension groups (see segment 3.8.14). We may therefore require that the realization of our subset

$$\tilde{\mathcal{E}}_n \subset \tilde{\mathcal{A}}_n$$

of near-extensions be linearly independent inside $\prod A(Z_{\mathfrak{p}})$; the precise statement is complicated by the fact that our realization map is itself merely an approximation. Computation of the inner products $\langle J, I\rangle_\epsilon$ for $J \in \mathcal{P}_n \cup \tilde{\mathcal{D}}_n$ reduces to computations in lower weights via the *Goncharov coproduct*, an explicit formula for the coproduct of two motivic iterated integrals due to Goncharov [2005]. For the remaining inner products, $\langle J, I\rangle_\epsilon$ with $J \in \tilde{\mathcal{E}}_n$, we use the realization algorithm to map the remaining part $I'$ of $I$ and $J$ into $\prod A(Z_{\mathfrak{p}})$ and compute there. Our requirement that $\mathrm{Re}_p\ \tilde{\mathcal{E}}_n$ be linearly independent ensures that the resulting system of linear equations will have a unique solution.

**1.10.** Given our abstract basis

$$\mathcal{A} = \mathcal{E} \cup \mathcal{P} \cup \mathcal{D}$$

of $A(Z^o)$, we obtain a set

$$\Sigma^o := \mathcal{E}^\vee$$

of free generators for the Lie algebra $\mathfrak{n}(Z^o)$. The set $\Phi$ of words in $\Sigma^o$ forms a basis for the universal enveloping algebra $\mathcal{U}(Z^o)$; its dual

$$\mathcal{F} \subset A(Z^o)$$

gives us a new basis, which plays nicely with the Hopf-algebra structure; we call such a basis a *shuffle basis*. We also have an exponential map

$$\exp^\sharp : U(Z^o) \xrightarrow{\sim} S^\bullet \mathfrak{n}(Z^o)^\vee$$

and we may compute the images

$$\exp^\sharp(f_w)$$

of the elements $f_w$ of $\mathcal{F}$ as Lie-words in $\Sigma^o$. We do not endeavor to carry this out explicitly here, contenting ourselves with the observation that the procedure is in an elementary sense algorithmic.

More interesting is the need to compare the two bases $\mathcal{A}$ and $\mathcal{F}$ of $A(Z^o)$. In segment 3.9, we approximate such a comparison by using our imperfect, concrete basis $\tilde{\mathcal{A}}$ to construct a near-shuffle basis $\tilde{\mathcal{F}}$, and by computing the associated change-of-basis matrix to given $p$-adic precision. This is our change of basis algorithm. An example is worked out in segment 7.6.3 of [Dan-Cohen and Wewers 2016].

**1.11.** In terms of a set of generators $\Sigma^o$ for the unipotent motivic Galois group $U(Z^o)$, the problem of computing the image of $\mathfrak{ev}_{\mathrm{Everywhere}}$ becomes purely classical (if not quite formal); we make this precise in segment 4.1 below. We let $\mathfrak{n}^{\mathrm{PL}}$ denote the Lie algebra of the polylogarithmic quotient $U^{\mathrm{PL}}$ of the

unipotent fundamental group of $X$, and we let $\mathfrak{n}(\Sigma^o)$ denote the free pronilpotent Lie algebra on $\Sigma^o$. The result is given as a finite family

$$\{F_i^{\mathrm{abs}}\}_i$$

of elements of

$$S^\bullet(\mathfrak{n}^{\mathrm{PL}}_{\geq -n} \times \mathfrak{n}(\Sigma^o)_{\geq -n})^\vee.$$

**1.12.** The unipotent fundamental group $U(X)$ at the tangent vector $1_0$ is canonically free prounipotent on two generators $e_0$, $e_1$ corresponding to monodromy about the punctures 0 and 1, respectively. As such, its coordinate ring $\mathcal{O}(U(X))$ possesses a canonical vector space basis $\{\mathrm{Li}_\omega\}_\omega$ where $\omega$ ranges over words in the two generators. We abbreviate words in $e_0$, $e_1$ by words in 0, 1. The polylogarithmic quotient

$$U(X) \twoheadrightarrow U^{\mathrm{PL}}$$

corresponds to the subalgebra generated by elements

$$\log := \mathrm{Li}_0, \quad \mathrm{Li}_1 := \mathrm{Li}_1, \quad \mathrm{Li}_2 := \mathrm{Li}_{10}, \quad \mathrm{Li}_3 := \mathrm{Li}_{100}, \quad \dots$$

In segment 4.2 we use the change-of-basis matrix of segment 3.9 to convert the elements $F_i^{\mathrm{abs}}$ into functions on

$$U(Z^o) \times U^{\mathrm{PL}}_{\geq -n}$$

given as elements $\tilde{F}_i$ of the polynomial ring

$$\mathbb{Q}[\tilde{\mathcal{E}} \cup \mathcal{P}, \log, \mathrm{Li}_1, \dots, \mathrm{Li}_n].$$

There is a natural map

$$\mathbb{Q}[\tilde{\mathcal{E}} \cup \mathcal{P}, \log, \mathrm{Li}_1, \dots, \mathrm{Li}_n] \to \mathrm{Col}(X(Z_{\mathfrak{p}}))$$

to the ring of Coleman functions, which we use to obtain the hoped-for family

$$\{\tilde{F}_i^{\mathfrak{p}}\}_i$$

of Coleman functions. Having completed the construction of the "Chabauty–Kim-loci" algorithm, $\mathcal{A}_{\mathrm{Loci}}$, we prove our main theorem in segment 4.2.6.

**1.13.** In Section 5 we use Newton polygons to bound the number of roots in small neighborhoods, and in Section 6 we review unpublished joint work with Andre Chatzistamatiou devoted to the computation of $p$-adic iterated integrals, on which we've already relied at several points.

We're now ready to construct the equation-solving algorithm of Theorem 1.2.1. Joint work with David Corwin [Corwin and Dan-Cohen 2018a] demonstrates the need for symmetrization with respect to the $S_3$ action on $X$. We set

$$X(Z_{\mathfrak{p}})_n^{S_3} := \bigcap_{\sigma \in S_3} \sigma(X(Z_{\mathfrak{p}})_n)$$

(see segment 2.1.3 for the precise definition). We search for points to obtain a gradually increasing list

$$X(Z)_n \subset X(Z).$$

At the same time we construct Coleman functions $\tilde{F}_i^{\mathfrak{p}}$ vanishing on $X(Z_{\mathfrak{p}})_n^{S_3}$ and use those to obtain a gradually decreasing union of neighborhoods. Thus, roughly speaking, $X(Z)$ is sandwiched

$$X(Z)_n \subset X(Z) \subset X(Z_{\mathfrak{p}})_n^{S_3}$$

with $X(Z)_n$ gradually increasing while $X(Z_{\mathfrak{p}})_n^{S_3}$ gradually decreases. We stop when the two sides meet. This concludes the construction of our equation-solving algorithm $\mathcal{A}_{\mathrm{ES}}$, and allows us to state and prove the equation-solving theorem (segment 7.2).

**1.14.** In Section 8 we generalize Theorem 1.3.1 to allow arbitrary open integer schemes. Essentially the only difference is that one is forced to replace $X(Z_{\mathfrak{p}})$ with the product $\prod_{\mathfrak{p}\,|\,p} X(Z_{\mathfrak{p}})$. We are unable at this point, however, to obtain an equation-solving algorithm in this generality: to do so we would have to study solutions to systems of locally analytic functions on higher-dimensional spaces.

**1.15.** ***Near-term goals.***

**1.15.1.** *Algorithmic precision.* The algorithmic constructions we make in this installment are precise by mathematical standards, but not by algorithmic standards, which are far more stringent. For instance, we do not attempt to make our $\epsilon$'s precise: any function of $\epsilon$ which is algorithmically computable, and which goes to zero with $\epsilon$, is again denoted by $\epsilon$ — we refer to this as an *admissible change in* $\epsilon$. Such imprecision is common in pure math, but useless for applications. Before going to Sage, we will of course have to compute exact levels of accumulated error as we make our approximations.

We also make no attempt to make our algorithm efficient: whenever we have a countable set, we don't hesitate to search through it arbitrarily. In fact, the problem of making our algorithm efficient interacts with significant, interesting problems of pure math; these include formulating explicit forms of Zagier's conjecture (available so far in only very special cases), and more precise versions of Goncharov's conjecture, at least with respect to ramification. Avoiding redundancy in our search through the set of iterated integrals is another interesting problem. Indeed, as Francis Brown has pointed out, as long as we limit ourselves to working with the polylogarithmic quotient, constructing a basis for all of $A(Z)_{\leq n}$ is huge overkill: any $\mathbb{G}_m$-equivariant map

$$U(Z) \to U(X)^{\mathrm{PL}}_{\geq -n}$$

must factor through a small quotient of $U(Z)$ (easily computed in terms of abstract coordinates). A careful study of this quotient should yield a conjecture which is both much weaker than Goncharov's conjecture, and much more efficient for us.[5]

[5]We alert the reader to the forthcoming works *The Goncharov quotient in computational Chabauty–Kim theory I and II* by the author and David Corwin in which this is carried out for open subschemes of Spec $\mathbb{Z}$ and new numerical results are obtained.

With regard to the endeavor to produce actual code, we would expect to push the computational boundary gradually, starting with very special cases in which the conjectures of Zagier and Goncharov are relatively well understood.

**1.15.2.** *Comparison with Brown's method.* In the spring of 2014 I visited Francis Brown at the IHES in order to discuss the previous installment in this series [Dan-Cohen and Wewers 2016]. Our meeting was inspiring and reassuring and helped me in developing the algorithm presented below.

Since then, Brown [2017] has made his own contribution to the subject in which he develops a method (or a kind of blueprint) for constructing *many* polylogarithmic functions on the $p$-adic points of $\mathbb{P}^1 \setminus \{0, 1, \infty\}$ over an open subscheme of $\operatorname{Spec} \mathbb{Z}$ which vanish on integral points, at least when there are *enough* integral points. His idea is that if Goncharov's conjecture is false, there should actually be *more* such functions available, increasing the chances of isolating the set of integral points. His assumption that $X$ has many points replaces our reliance on Goncharov's conjecture for halting, and one of his goals, which he achieves in several examples, is to circumvent our construction of a basis of the mixed Tate–Hopf algebra $A(Z)$. A particularly satisfying outcome is a more economic and aesthetic construction of the polylogarithmic function constructed in [Dan-Cohen and Wewers 2016].[6]

Work remains to be done in comparing Brown's construction with ours, and hopefully, in harnessing the power of both approaches to construct more efficient, and more enlightening, algorithms.

**1.15.3.** *Beyond totally real fields.* Most of the work completed here applies to arbitrary open integer schemes which obey *Kim vs. Hasse*. For the final application, however, we are limited to the totally real case. As mentioned above, in order to go further, we would have to develop methods for computing solutions to systems of polylogarithmic functions in higher dimensions.

**1.15.4.** *Beyond the polylogarithmic quotient.* Beyond the polylogarithmic quotient, the motivic Selmer variety

$$H^1(G(Z), U(X)_{\geq -n})$$

is still canonically isomorphic to the space

$$Z^1(U(Z), U(X)_{\geq -n})^{\mathbb{G}_m}$$

of $\mathbb{G}_m$-equivariant cocycles. Explicit computation of this space is complicated however by the fact that the action of $U(Z)$ on $U(X)$ is highly nontrivial. This task is urgent for at least two reasons. Obviously, replacing the polylogarithmic quotient with the full unipotent fundamental group in our current algorithm would enable us to weaken our version of Kim's conjecture. More interesting, perhaps, would be the

[6] I should point out, however, that the purpose of the present work is somewhat different from Brown's work. Our goal here is to construct an actual algorithm, complete with precise criteria for halting. Moreover, its output should consist of functions that vanish not only on integral points, but also on the (a priori larger) loci defined by Kim's general theory. This ensures that the result may be used to produce numerical evidence for Kim's conjecture on the one hand, and provides an example of how Kim's framework for studying integral points may be made fully explicit and algorithmic on the other hand.

possibility of going beyond our three punctures $\{0, 1, \infty\}$ to more general punctures, including possibly punctures that are not rational over the base-field in the context of mixed Artin–Tate motives.

## 2. Conjectures and theorems

We begin with a brief review of background material (unipotent fundamental group, Kim's conjecture, motivic iterated integrals, filtered $\phi$ iterated integrals, $p$-adic periods). For a more detailed exposition, tailored specifically to our applications, we refer the reader to Dan-Cohen and Wewers [2016].

### 2.1. *A motivic variant of Kim's construction.*

**2.1.1.** The prounipotent completion of the fundamental group of $X$ has a motivic precursor, known as the *unipotent fundamental group*, a unipotent group object of the category of mixed Tate motives constructed by Deligne and Goncharov [2005] whose various realizations had previously been studied by Deligne [1989]. We use the tangent vector $1_0$ as base-point, and denote the resulting group simply by $U(X)$; see Section 15 of [loc. cit.] for the use of tangential base-points in the various realizations. The unipotent fundamental group of $\mathbb{G}_m$ is equal to (the covariant total space of) $\mathbb{Q}(1)$. The natural inclusion

$$\mathbb{P}^1 \setminus \{0, 1, \infty\} \hookrightarrow \mathbb{G}_m$$

induces a surjection of unipotent fundamental groups

$$U(X) \twoheadrightarrow \mathbb{Q}(1).$$

Let $N$ denote its kernel. Then according to Deligne [loc. cit., Section 16], the Lie algebra of

$$U(X)^{\mathrm{PL}} := U(X)/[N, N]$$

is canonically a semidirect product

$$\mathfrak{n}(X)^{\mathrm{PL}} = \mathbb{Q}(1) \ltimes \prod_{i=1}^{\infty} \mathbb{Q}(i).$$

We write $\mathfrak{n}(X)^{\mathrm{PL}}_{\geq -n}$ for the quotient

$$\mathbb{Q}(1) \ltimes \prod_{i=1}^{n} \mathbb{Q}(i),$$

of $\mathfrak{n}(X)^{\mathrm{PL}}$, and $U^{\mathrm{PL}}_{\geq -n}$ for the corresponding quotient of $U(X)$. There are also associated quotients $({}_b P_{1_0})^{\mathrm{PL}}_{\geq -n}$ of the path torsors ${}_b P_{1_0}$ obtained by pushing out along the quotient map of

$$U \twoheadrightarrow U^{\mathrm{PL}}_{\geq -n}.$$

All this holds over $\operatorname{Spec} \mathbb{Z}$.

**2.1.2.** Let $Z \subset \operatorname{Spec} \mathcal{O}_K$ be an open integer scheme. Sending a point $b \in X(Z)$ to the torsor of *polylogarithmic paths* ${}_b P_{1_0})^{\mathrm{PL}}_{\geq -n}$ defines a map

$$\kappa : X(Z) \to H^1(U^{\mathrm{PL}}_{\geq -n})(\mathbb{Q})$$

to the set of rational points of a finite-type affine $\mathbb{Q}$-scheme $H^1(U^{\mathrm{PL}}_{\geq -n})$ which parametrizes such torsors, the *polylogarithmic Selmer variety*; see Section 2.3.2 below for a precise definition.

There is also a local $p$-adic version. We fix a closed point $\mathfrak{p}$ of $Z$ which we assume to be totally split for simplicity. We write $\mathcal{O}_\mathfrak{p}$ instead of $\mathbb{Z}_p$ when we wish to emphasize the $\mathcal{O}_Z$-algebra structure, and similarly for $K_\mathfrak{p} = \mathbb{Q}_p$. We denote $\operatorname{Spec} \mathcal{O}_\mathfrak{p}$ by $Z_\mathfrak{p}$. We obtain a map

$$\kappa_\mathfrak{p} : X(\mathcal{O}_\mathfrak{p}) \to H^1(U^{\mathrm{PL},F\phi}_n)(\mathbb{Q}_p)$$

to the set of $\mathbb{Q}_p$-points of the *filtered-$\phi$ polylogarithmic Selmer variety* $H^1(U^{\mathrm{PL},F\phi}_n)$. The set $X(\mathcal{O}_\mathfrak{p})$ may be viewed as the set of $\mathbb{Q}_p$-points of the rigid analytic space (for instance) obtained from $\mathbb{P}^1_{\mathbb{Q}_p}$ by removing the residue disks about 0, 1, and $\infty$. Viewed this way, the map $\kappa_\mathfrak{p}$ is locally analytic but not rigid analytic; rather, it is given in coordinates by Coleman functions.

Connecting the filtered-$\phi$ and motivic versions is a map of $\mathbb{Q}_p$-schemes

$$\mathfrak{R}_\mathfrak{p} : \mathbb{Q}_p \otimes H^1(U^{\mathrm{PL}}_{\geq -n}) \to H^1(U^{\mathrm{PL},F\phi}_{\geq -n})$$

induced by $p$-adic de Rham realization, which forms a commuting square

$$\begin{array}{ccc} X(Z) & \longrightarrow & X(\mathcal{O}_\mathfrak{p}) \\ \downarrow & & \downarrow{\scriptstyle \kappa_\mathfrak{p}} \\ H^1(U^{\mathrm{PL}}_{\geq -n})(\mathbb{Q}_p) & \xrightarrow[\mathfrak{R}_\mathfrak{p}]{} & H^1(U^{\mathrm{PL},F\phi}_{\geq -n})(\mathbb{Q}_p). \end{array}$$

**2.1.3.** The following conjecture was formulated jointly with David Corwin in [Corwin and Dan-Cohen 2018a]. Let $\mathrm{Col}(X(\mathcal{O}_\mathfrak{p}))$ denote the ring of Coleman functions and consider the induced maps of $\mathbb{Q}_p$-algebras:

$$\begin{array}{ccc} & & \mathrm{Col}(X(\mathcal{O}_\mathfrak{p})) \\ & & \uparrow{\scriptstyle \kappa^\sharp_\mathfrak{p}} \\ \mathbb{Q}_p \otimes \mathcal{O}(H^1(U^{\mathrm{PL}}_{\geq -n})) & \xleftarrow[\mathfrak{R}^\sharp_\mathfrak{p}]{} & \mathcal{O}(H^1(U^{\mathrm{PL},F\phi}_{\geq -n})) \end{array}$$

There is a natural action of the symmetric group $S_3$ on $\mathbb{P}^1 \setminus \{0, 1, \infty\}$, hence also on $\mathrm{Col}(X(\mathcal{O}_\mathfrak{p}))$. We let $\kappa^\sharp_\mathfrak{p}(\ker \mathfrak{R}^\sharp_\mathfrak{p})$ denote the ideal of $\mathrm{Col}(X(\mathcal{O}_\mathfrak{p}))$ generated by the image of $\ker \mathfrak{R}^\sharp_\mathfrak{p}$ and we let

$$\kappa^\sharp_\mathfrak{p}(\ker \mathfrak{R}^\sharp_\mathfrak{p}) S_3$$

denote the ideal generated by its orbit. In down to earth terms, this means that we close the set of generators $\{F_i(z)\}_i$ of the smaller ideal under the two operations

$$F_i \mapsto F_i(1-z), \quad \text{and} \quad F_i \mapsto F_i\big(\tfrac{1}{z}\big). \tag{$*$}$$

We first define the *polylogarithmic Chabauty–Kim locus at level n*

$$X(\mathcal{O}_\mathfrak{p})_n \subset X(\mathcal{O}_\mathfrak{p})$$

to be the vanishing locus (not a priori reduced) of the smaller ideal $\kappa_\mathfrak{p}^\sharp(\ker \mathfrak{R}_\mathfrak{p}^\sharp)$. The polylogarithmic Chabauty–Kim loci form a nested sequence

$$X(\mathcal{O}_\mathfrak{p}) \supset X(\mathcal{O}_\mathfrak{p})_1 \supset X(\mathcal{O}_\mathfrak{p})_2 \supset \cdots \supset X(Z).$$

We note the following theorem, which is a direct consequence of the results of Kim [2012] via Soulé's étale regulator isomorphism [1981].

**Theorem** (Kim). *Suppose $Z$ is a totally real open integer scheme and let $\mathfrak{p} \in Z$ be any prime. Then for $n$ sufficiently large,* $\operatorname{Im}\mathfrak{R}_\mathfrak{p}$ *is contained in a subscheme of $H^1(U^{\mathrm{PL},F\phi}_{\geq -n})$ of strictly lower dimension.*

Recall from Kim [2009] that the map $\kappa_\mathfrak{p}$ has dense image. So as soon as we have a nonzero function on $H^1(U^{\mathrm{PL},F\phi}_{\geq -n})$ vanishing on $\operatorname{Im}\mathfrak{R}_\mathfrak{p}$, the associated locus will be finite.

**Corollary** (Kim). *Suppose $Z$ is a totally real open integer scheme, and assume $\mathfrak{p} \in Z$ is totally split. Then for $n$ sufficiently large, the associated polylogarithmic Chabauty–Kim locus $X(\mathcal{O}_\mathfrak{p})$ is finite.*

We define the *symmetrized polylogarithmic Chabauty–Kim locus at level n*

$$X(\mathcal{O}_\mathfrak{p})_n^{S_3} \subset X(\mathcal{O}_\mathfrak{p})$$

to be the vanishing locus (not a priori reduced) of the ideal $\kappa_\mathfrak{p}^\sharp(\ker \mathfrak{R}_\mathfrak{p}^\sharp)S_3$. Stretching Kim's conjecture from [Balakrishnan et al. 2018] somewhat, we propose the following.

**Conjecture 2.1.4** (Convergence of polylogarithmic loci, joint with David Corwin). *Let $Z$ be a totally real open integer scheme, and $\mathfrak{p} \in Z$ a totally split prime. View $X(Z)$ as a locally analytic space over $\mathbb{Q}_p$ with reduced structure. Then for $n$ sufficiently large, the associated polylogarithmic Chabauty–Kim locus satisfies*

$$X(\mathcal{O}_\mathfrak{p})_n^{S_3} = X(Z).$$

**Remark 2.1.5.** The generalization from the rational to the totally real case should be harmless. By restricting attention to the polylogarithmic quotient, however, we are relying on a proper strengthening of Kim's conjecture. Nevertheless, since the codimension of $\operatorname{Im}\mathfrak{R}_\mathfrak{p}$ goes to infinity already for the polylogarithmic quotient, much of the motivation for Kim's conjecture does hold for the polylogarithmic quotient; see [Corwin and Dan-Cohen 2018a] for a discussion of the role played by the $S_3$-orbit. Finally, our interpretation of $X(\mathcal{O}_\mathfrak{p})_n$ as a potentially nonreduced space, implying that there should be no double roots for $n$ sufficiently large, is not discussed explicitly in [Balakrishnan et al. 2018].

### 2.2. *Iterated integrals, p-adic periods, statement of arithmetic conjectures.*

**2.2.1.** We begin by reviewing those properties of mixed Tate motives that we use. This material may be found, for instance, in [Deligne and Goncharov 2005]. We continue to work with an open integer scheme $Z$. Let $\boldsymbol{MT}(Z)$ denote the category of (unramified) mixed Tate motives over $Z$ with $\mathbb{Q}$-coefficients. The category $\boldsymbol{MT}(Z)$ is $\mathbb{Q}$-Tannakian. It has a special object $\mathbb{Q}(1)$ of rank 1. Every simple object is isomorphic to a unique $\mathbb{Q}(n) := \mathbb{Q}(1)^{\otimes n}$. Each object is equipped with an increasing filtration $W$, *the Weight filtration*. (Since all the weights that occur are even, we also work with *half-weights*, which are half the usual weights. These will be denoted by subscripts.) The functor

$$\boldsymbol{MT}(Z) \to \mathrm{Vect}(\mathbb{Q})$$

sending

$$E \mapsto \bigoplus \mathrm{Hom}(\mathbb{Q}(i), \mathrm{gr}^W_{-2i} E)$$

is a $\mathbb{Q}$-valued fiber functor, with associated group of the form

$$G(Z) = U(Z) \rtimes \mathbb{G}_m$$

with $U(Z)$ free prounipotent. From generalities of mixed Tate categories, we have canonical isomorphisms

$$U(Z)^{\mathrm{ab}} = \bigoplus_{i \geq 1} \mathrm{Ext}^1_Z(\mathbb{Q}(0), \mathbb{Q}(n))^\vee.$$

We have (highly nontrivial) canonical isomorphisms

$$K^{(n)}_{2n-1}(Z) \xrightarrow{\sim} \mathrm{Ext}^1_Z(\mathbb{Q}(0), \mathbb{Q}(n)),$$

and a computation of the dimensions of these $K$-groups via real-analytic methods due to Borel [1953; 1977]:

$$\dim K^{(n)}_{2n-1}(Z) = \begin{cases} r_1 + r_2 & \text{for } n \text{ odd } \geq 3, \\ r_2 & \text{for } n \text{ even } \geq 2, \end{cases}$$

where $r_1$ (resp. $r_2$) denotes the number of real (resp. complex) places.

We let $\mathfrak{n}(Z)$ denote the Lie algebra of $U(Z)$, $\mathcal{U}(Z)$ its completed universal enveloping algebra, and $A(Z)$ the coordinate ring of $U(Z)$. Recall that the natural map

$$\mathcal{U}(Z)^\vee \to A(Z)$$

is an isomorphism of $\mathbb{Q}$-vector spaces.

**2.2.2.** Our discussion of iterated integrals applies to the complement in $\mathbb{P}^1_Z$ of any divisor $\boldsymbol{D}$ which is a union of sections of

$$\mathbb{P}^1_Z \to Z$$

and is étale over $Z$; we continue to use the letter $X$ which now denotes $\mathbb{P}^1_Z \setminus \boldsymbol{D}$ and we refer to the components of $\boldsymbol{D}$ as *punctures*. We assume $\infty \in \boldsymbol{D}$. We say that a section of a vector bundle is *nowhere*

*vanishing* if its image in every closed fiber is nonzero. We define a *base-point* to be either an integral point or a nowhere vanishing section of the normal bundle to one of the punctures. If $a$ is a base-point, we denote the unipotent fundamental group of $X$ at $a$ by $U_a(X)$. The unipotent fundamental group may be thought of as a prounipotent group object of $\boldsymbol{MT}(Z)$, or, after applying the canonical fiber functor, as a prounipotent $\mathbb{Q}$-group equipped with an action of $G(Z)$, and we do not distinguish between these two points of view when we see no cause for confusion.

If $b$ is a second base-point, we denote the unipotent path torsor by ${}_bP_a$; the latter may be thought of internally as a torsor-object of $\boldsymbol{MT}(Z)$ or externally as a $G(Z)$-equivariant torsor.

**2.2.3.** After forgetting the $G(Z)$-action, each unipotent path torsor ${}_bP_a$ is trivialized by a special $\mathbb{Q}$-rational path

$$ {}_b p_a^{\mathrm{dR}} \in {}_bP_a(\mathbb{Q}), $$

and the fundamental group $U_a(X)$ is free on the set of logarithmic vector fields dual to the 1-forms

$$ \omega_c = \frac{dt}{t-c} $$

for $c$ a component of $\boldsymbol{D}_f := \boldsymbol{D}\backslash\infty$; this is proved by Deligne [1989] when $K=\mathbb{Q}$ and by Goncharov [2005] in general.[7][8] If $\omega = (\omega^1, \ldots, \omega^r)$ is a sequence of such differential forms, we let $f_\omega$ denote the associated function (see Section 2 of [Dan-Cohen and Wewers 2016] for generalities on free prounipotent groups).

We say that the datum $(a;\omega;b)$ is *combinatorially unramified* if the associated reduced divisors are étale over $Z$.[9] Being combinatorially unramified has the effect that the entire path bimodule $\mathcal{U}_bP_a$ (i.e., the bimodule over the completed universal enveloping algebras at $a$ and $b$) is unramified over $Z$.

**2.2.4.** Following Goncharov [2005], we define $I_a^b(\omega)$ to be the composite

$$ U(Z) \xrightarrow{o({}_b p_a^{\mathrm{dR}})} {}_bP_a(X) \xrightarrow{\sim} U_a(X) \xrightarrow{f_\omega} \mathbb{A}^1_{\mathbb{Q}}. $$

Here $o({}_b p_a^{\mathrm{dR}})$ denotes the orbit map associated to the rational point ${}_b p_a^{\mathrm{dR}}$. If $A(Z)$ denotes the graded hopf algebra $\mathcal{O}(U(Z))$ then $I_a^b(\omega)$ belongs to $A(Z)_r$. We refer to these elements as *(combinatorially unramified) unipotent iterated integrals*. Among the unipotent iterated integrals are the *classical unipotent polylogarithms*

$$ \mathrm{Li}^U_{n+1}(t) := I^t_{1_0}(0^{\cdot n}, 1) $$

[7] It is quite crucial that we work rationally here, since we will be using motivic iterated integrals to construct generators of the Hopf algebra $A(Z) = \mathcal{U}(Z)^\vee$ as a $\mathbb{Q}$-algebra.

[8] After tensorization with $K$, the canonical fiber functor on $\boldsymbol{MT}(Z)$ becomes canonically isomorphic to the de Rham fiber functor. A fact, which is perhaps underemphasized in the literature, however, is that the usual Tannakian interpretation of $U_a(X)$ in terms of unipotent connections is unavailable over the rationals unless $K=\mathbb{Q}$. Thus, our ${}_b p_a^{\mathrm{dR}}$ is a *rational form* of Deligne's canonical de Rham path.

[9] If $a$ is a base-point, we let $a_0$ denote the *location* of $a$: $a_0 = a$ if $a$ is an integral point, otherwise $a$ is a tangent vector *at* $a_0$. We assume that the datum $(a;\omega;b)$ behaves nicely over $Z$ in an obvious sense, which requires several cases to state precisely: in all cases we assume the reduced divisor associated to $\omega$, $a_0$, $b_0$, and $\infty$ is étale over $Z$; if $a$ is a tangent vector and $a_0 \neq b_0$, we assume the reduced divisor which supports $a+0+\infty$ on $\mathbb{P}^1$ is étale over $Z$; if $a$, $b$ are both tangent vectors at the same point $a_0 = b_0$, we assume similarly that the support of $a+b+0+\infty$ is étale over $Z$.

(where the comma is used as a typographical pun to denote the concatenation product) and their single valued cousins $\mathrm{Li}_n^{U,sv}(t)$, for which we refer the reader to Brown [2013]. In terms of these objects, we may state the conjectures of Zagier and Goncharov as follows.

**Conjecture 2.2.5** (Zagier's conjecture). *For each $n \geq 2$, the motivic Ext group*

$$E_n := \mathrm{Ext}^1_K(\mathbb{Q}(0), \mathbb{Q}(n))$$

*is spanned by single valued unipotent $n$-logarithms $\mathrm{Li}_n^{U,sv}(t)$ with $t \in K$.*

**Remark 2.2.6.** We recall that Zagier's conjecture is known for $n = 2$ by Zagier [1991] and independently by work of Suslin [1987] and Bloch [2000], and for $n = 3$ by Goncharov [1994]. The case of $K$ cyclotomic was treated by Beilinson [1989] in an unpublished manuscript. The main algorithm we construct below could be greatly simplified in the cyclotomic case $K = \mathbb{Q}(\zeta_N)$, where a basis for $E_n$ is given explicitly by the elements $\mathrm{Li}_n^{U,sv}(\zeta_N^i)$ for $0 < i < N/2$. Conversely, away from the cyclotomic case, our algorithm is made complicated partly because of the lack of explicit constructions, even conjectural, of elements of $E_n$; away from the roots of unity, most $\mathrm{Li}_n^{U,sv}(t)$ are *not* contained in $E_n$.

**Conjecture 2.2.7** (Goncharov-exhaustion). *For any open integer scheme $Z$ and any $n \in \mathbb{N}$, there is an open subscheme $Z^o \subset Z$ such that $A(Z^o)_{\leq n}$ is spanned by combinatorially unramified unipotent iterated integrals over $Z^o$.*

**Remark 2.2.8.** Goncharov's conjecture [1995] implies that for every $Z$ and $n$ there is an open subscheme $Z^o \subset Z$ such that $A(Z)_{\leq n}$ (in place of $A(Z^o)_{\leq n}$) is spanned by *linear combinations of* combinatorially unramified unipotent iterated integrals over $Z^o$. This distinction between actual iterated integrals and linear combinations of iterated integrals is important, and our strengthening of the conjecture is, as far as I can tell, nontrivial and necessary for our purposes. The reason is that we are unable to check algorithmically if a given linear combination of combinatorially unramified iterated integrals in $A(Z^o)$ belongs to the subalgebra

$$A(Z) \subset A(Z^o).$$

Actually, our work here does provide an algorithm for checking if a given linear combination is $p$-adically close to $A(Z)$, but this algorithm is rather indirect. In outline, we first apply our *basis algorithm* to obtain an open subscheme $Z^o \subset Z$ and a *concrete* basis of $A(Z^o)$ (within the specified weight range) consisting of (actual!) combinatorially unramified unipotent iterated integrals compatible to within $\epsilon$ with the extension spaces. This gives rise to a set of generators of the Lie algebra, which in turn generate an *abstract* shuffle basis of $A(Z^o)$. We then compare the two bases to within $\epsilon$ using our *change of basis algorithm* and use the shuffle basis to identify $A(Z)$ inside $A(Z^o)$.

Thus, our conjecture represents a version of Goncharov's conjecture strengthened somewhat to include weak control over ramification. Better control over ramification would yield a faster algorithm. The case $Z = \mathrm{Spec}\,\mathbb{Z} \setminus \{2\}$, $n = \infty$ established by Deligne [2010], and the discussion of the case $n = 2$

in [Dan-Cohen and Wewers 2016] both support the belief that unipotent iterated integrals should be compatible with ramification, at least to the extent predicted by our wording of the conjecture.

**2.2.9.** Unipotent iterated integrals have a filtered $\phi$ variant at each prime $\mathfrak{p} \in Z$. We mention only a few key similarities and differences, referring the reader to [Dan-Cohen and Wewers 2016, Section 4] for details. As for mixed Tate motives, there is a Tannakian (in fact, mixed Tate) category of mixed Tate filtered $\phi$ modules, and an associated proalgebraic group of the form

$$G(\mathcal{O}_\mathfrak{p}) = \mathbb{G}_m \ltimes U(\mathcal{O}_\mathfrak{p})$$

with $U(\mathcal{O}_\mathfrak{p})$ free prounipotent, but now over $\mathbb{Q}_p$; we adopt our notation $(\mathfrak{n}(\mathcal{O}_\mathfrak{p}), \mathcal{U}(\mathcal{O}_\mathfrak{p}), A(\mathcal{O}_\mathfrak{p}))$ from the motivic case.[10] Unlike the motivic case, $U(\mathcal{O}_\mathfrak{p})$ possesses canonical generators $v_{\mathfrak{p},-1}, v_{\mathfrak{p},-2}, v_{\mathfrak{p},-3}, \ldots$, and an associated special $K_\mathfrak{p}$-valued point

$$u_\mathfrak{p} = \exp \sum_i v_{\mathfrak{p},i}$$

of $U(\mathcal{O}_\mathfrak{p})$. Note that the "generators" are not points of the group, but rather elements of the Lie algebra, regarded as Lie-like elements of the completed universal enveloping algebra — see the discussion of free prounipotent groups in Section 2 of [Dan-Cohen and Wewers 2016].

**2.2.10.** There is a morphism of unipotent groups

$$U(Z) \leftarrow U(Z_\mathfrak{p})$$

(linear over $\operatorname{Spec} \mathbb{Q} \leftarrow \operatorname{Spec} \mathbb{Q}_p$) induced by filtered $\phi$ realization. The composite

$$U(Z) \leftarrow U(Z_\mathfrak{p}) \xleftarrow{u_\mathfrak{p}} \operatorname{Spec} K_\mathfrak{p}$$

is the map denoted $I_{\mathrm{BC}}$ above; we refer to it as "Besser–Coleman integration". The associated map of rings

$$\mathrm{per}_\mathfrak{p} := I_{\mathrm{BC}}^\sharp : A(\mathcal{O}_\mathfrak{p}) \to K_\mathfrak{p}$$

is called the *p-adic period map*. If $\pi$ denotes the embedding of $K$ in $K_\mathfrak{p}$, then we have

$$I_a^b(\omega)(I_{\mathrm{BC}}) = \mathrm{per}_\mathfrak{p}(I_a^b(\omega)) = \int_{a^\pi}^{b^\pi} \omega^\pi,$$

a $p$-adic iterated integral in the sense of Coleman–Besser. (From this point of view, it's better to think of $I_a^b(\omega)$ as a "motivic iterated *integrand*": when we combine a *motivic iterated integrand* with *p-adic integration*, we obtain a *p-adic iterated integral*.) An algorithm for computing such integrals to arbitrary $p$-adic precision is constructed in [Dan-Cohen and Chatzistamatiou 2014]; as mentioned above, we review this unpublished work in Section 6 below. The following conjecture is stated for instance in Yamashita [2010].

[10] Our use of $\mathcal{O}_\mathfrak{p}$ (in place of $K_\mathfrak{p}$) in the notation expresses the fact that we're working with filtered $\phi$-modules as opposed to filtered $\phi, N$-modules.

**Conjecture 2.2.11** ($p$-Adic period conjecture). *Let $Z$ be an open integer scheme with fraction field $K$, and let $\mathfrak{p}$ be a closed point of $Z$. Then the $p$-adic period map*

$$\mathrm{per}_{\mathfrak{p}} : A(Z) \to K_{\mathfrak{p}}$$

*is injective.*

**2.2.12.** *Hasse principle for finite cohomology.* In addition to the semilinear injectivity of the period conjecture, we will also need a linear injectivity property which concerns the product of realization maps

$$\mathfrak{R}_p : \mathbb{Q}_p \otimes \mathrm{Ext}^1_{\mathcal{O}_K}(\mathbb{Q}(0), \mathbb{Q}(n)) \to \prod_{\mathfrak{p} \mid p} \mathrm{Ext}^1_{\mathcal{O}_{\mathfrak{p}}}(\mathbb{Q}_p(0), \mathbb{Q}_p(n))$$

for $n \geq 2$. We recast this in the language of finite Galois cohomology as follows. Let $S$ be the set of places of $K$ above $p$ and $\infty$. Let $G_S$ denote the Galois group of the maximal extension of $K$ which is unramified outside of $S$, and for $v$ a place of $K$, let $G_v$ denote the total Galois group of the local field $K_v$. Following Bloch and Kato [1990], we write $H^i_f$ for the space of cohomology classes that are crystalline at all primes above $p$. By the $p$-adic regulator isomorphisms of Soulé [1981]

$$\mathbb{Q}_p \otimes \mathrm{Ext}^1_{\mathcal{O}_K[1/p]}(\mathbb{Q}(0), \mathbb{Q}(n)) \xrightarrow{\sim} H^1(G_S, \mathbb{Q}_p(n)),$$

the injectivity of $\mathfrak{R}_p$ is equivalent to the following condition on a number field $K$, a prime $p$ of $\mathbb{Z}$ and an integer $n \geq 2$.

**Condition 2.2.13.** The map

$$\mathrm{loc}_p : H^1_f(G_S, \mathbb{Q}_p(n)) \to \prod_{\mathfrak{p} \mid p} H^1_f(G_{\mathfrak{p}}, \mathbb{Q}_p(n))$$

is injective.

Let $Z$ be a totally real open integer scheme with function field $K$ and assume the corresponding polylogarithmic Chabauty–Kim loci converge at $n$. We say that $Z$ *obeys Kim vs. Hasse* if the above injectivity holds at levels $n' \leq n$.

In fact, an anonymous referee has pointed out that Condition 2.2.13 follows from a conjecture due to Jannsen. Conjecture 1 of [Jannsen 1989] (applied, in the notation of that article, to $X = \mathrm{Spec}\, \mathcal{O}_K[1/p]$) says that

$$H^2(G_S, \mathbb{Q}_p(n)) = 0$$

for $n < 0$. By the Poitou–Tate exact sequence

$$\cdots \to H^2(G_S, M^\vee(1))^\vee \to H^1(G_S, M) \to \bigoplus_{v \in S} H^1(G_v, M) \to \cdots$$

applied to $M = \mathbb{Q}_p(n)$, we find that the map

$$H^1(G_S, \mathbb{Q}_p(n)) \to \bigoplus_{v \in S} H^1(G_v, \mathbb{Q}_p(n))$$

is injective whenever $n \geq 2$.

For $v$ real and $p$ odd, we have

$$H^1(G_v, \mathbb{Q}_p(n)) = 0.$$

Indeed, if $C$ is a finite cyclic group with generator $\sigma$, and if we consider the elements $1-\sigma$, $N := \sum_{\tau\in C} \tau$ of the group algebra $\mathbb{Z}[C]$, then for any $\mathbb{Z}[C]$-module $A$, the sequence

$$0 \to A \xrightarrow{\sigma-1} A \xrightarrow{N} A \xrightarrow{\sigma-1} A \xrightarrow{N} \cdots,$$

in which the first $A$ is in degree zero, forms a complex $A^\bullet$ and

$$H^i(C, A) = H^i A^\bullet.$$

When $C$ has order 2, we have $N = \sigma + 1$. Applying this to our situation, we have

$$H^1(G_v, \mu_{p^r}^{\otimes n}) = H^1(\mathbb{Z}/(2), (\mathbb{Z}/(p^r))^{\otimes n})$$

computed by the complex

$$0 \to \mathbb{Z}/(p^r) \xrightarrow{(-1)^n-1} \mathbb{Z}/(p^r) \xrightarrow{(-1)^n+1} \mathbb{Z}/(p^r) \to \cdots$$

in which isomorphisms alternate with multiplication by $\pm 2$ depending on the parity of $n$. Either way, all cohomologies above degree 0 vanish.

Consequently, the map

$$H^1(G_S, \mathbb{Q}_p(n)) \to \bigoplus_{\mathfrak{p}\mid p} H^1(G_\mathfrak{p}, \mathbb{Q}_p(n))$$

is injective. Condition 2.2.13 follows by restricting this map to finite cohomology spaces.

As a final remark, let us note that this injectivity is related to the nonvanishing of certain $p$-adic $L$-values; see Theorem 4.2.1 of Perrin–Riou [1994].

### 2.3. *Outline of algorithm.*

**2.3.1.** Our main construction is an algorithm, which we denote by $\mathcal{A}_{\mathrm{Loci}}$, which takes as input an open integer scheme $Z$, a prime $p$ of $\mathbb{Z}$ over which $Z$ is totally split, a natural number $n$, and an $\epsilon$, and returns an open subscheme $Z^o$ of $Z$, an algebra basis $\tilde{\mathcal{B}}$ of the polynomial ring $A(Z^o)_{\le n}$, and a family $\{\tilde{F}_i\}_i$ of elements of the polynomial ring

$$\mathbb{Q}[\tilde{\mathcal{B}}, \log, \mathrm{Li}_1, \ldots, \mathrm{Li}_n].$$

**2.3.2.** In terms of the category of mixed Tate motives $\boldsymbol{MT}(Z)$ and its Tannakian fundamental group $G(Z)$ discussed in Section 2.2.1, the polylogarithmic Selmer variety of Section 2.1.2 is characterized by the functor of $\mathbb{Q}$-algebras

$$R \mapsto H^1(G(Z)_R, U^{\mathrm{PL}}_{\ge -n,R}).$$

The proof of Proposition 2 of Kim [2005] applies mutatis mutandis to show that this functor is representable by a finite-type affine $\mathbb{Q}$-scheme, which in this case is in fact isomorphic to affine space.

**2.3.3.** According to [Dan-Cohen and Wewers 2016, Section 5.2], we have

$$H^1(G(Z), U(X)^{\mathrm{PL}}_{\geq -n}) = Z^1(U(Z), U(X)^{\mathrm{PL}}_{\geq -n})^{\mathbb{G}_m} = \mathrm{Hom}(U(Z), U(X)^{\mathrm{PL}}_{\geq -n})^{\mathbb{G}_m},$$

the space of $\mathbb{G}_m$-equivariant homomorphisms, and similarly for the filtered $\phi$ version over $Z_{\mathfrak{p}}$. Moreover, in the latter case, evaluation at $u_{\mathfrak{p}}$ induces an isomorphism

$$\mathfrak{ev}_{u_{\mathfrak{p}}} : \mathrm{Hom}(U(Z_{\mathfrak{p}}), U(X_{\mathfrak{p}})^{\mathrm{PL}}_{\geq -n})^{\mathbb{G}_m} \xrightarrow{\sim} U(X_{\mathfrak{p}})^{\mathrm{PL}}_{\geq -n} = \mathbb{Q}_p \otimes U(X)^{\mathrm{PL}}_{\geq -n}.$$

The composite map

$$\mathbb{Q}_p \otimes \mathrm{Hom}(U(Z), U(X)^{\mathrm{PL}}_{\geq -n})^{\mathbb{G}_m} \to \mathrm{Hom}(U(Z_{\mathfrak{p}}), U(X_{\mathfrak{p}})^{\mathrm{PL}}_{\geq -n})^{\mathbb{G}_m} \xrightarrow{\sim} \mathbb{Q}_p \otimes U(X)^{\mathrm{PL}}_{\geq -n}$$

is given by evaluation at the pullback $I_{\mathrm{BC}}$ of $u_{\mathfrak{p}}$ to $U(Z)$. As explained in the introduction, in order to compute its scheme-theoretic image, we first put this evaluation map inside the universal family of evaluation maps $\mathfrak{ev} = \mathfrak{ev}_{\mathrm{Everywhere}}$ :

$$\mathrm{Hom}^{\mathbb{G}_m}(U(Z), U(X)^{\mathrm{PL}}_{\geq -n}) \times U(Z^o) \to U(X)^{\mathrm{PL}}_{\geq -n} \times U(Z^o)$$

pulled back along

$$U(Z^o) \twoheadrightarrow U(Z).$$

If we fix arbitrary generators of $U(Z^o)$, these give rise to coordinates on $A(Z^o)$, which we refer to as *abstract shuffle-coordinates*. In terms of these, the computation is purely classical. We must then however switch to coordinates whose image under the period map can be computed, that is, to *concrete coordinates* given by unipotent iterated integrals. As explained in the introduction, the heart of our algorithm constructs such coordinates, as well as an approximate change-of-basis matrix which relates a judicious choice of abstract shuffle-coordinates to our concrete coordinates. This key step is inspired by the work of Francis Brown [2012].

**2.4.** ***Statement of main theorem.*** For each prime $\mathfrak{p}$ lying above $p$, the $\mathfrak{p}$-adic period map extends in an obvious way to a map

$$\mathbb{Q}[\tilde{\mathcal{B}}, \log, \mathrm{Li}_1, \ldots, \mathrm{Li}_n] \to \mathrm{Col}(X(\mathcal{O}_{\mathfrak{p}}))$$

to the ring of Coleman functions; denote the image of the element $\tilde{F}_i$ from segment 2.3.1 by $\tilde{F}_i^{\mathfrak{p}}$.

**Theorem 2.4.1.** *Let $Z$ be an open integer scheme, $\mathfrak{p} \in Z$ a totally split prime, $p$ the image of $\mathfrak{p}$ in* $\mathrm{Spec}\,\mathbb{Z}$, *$n$ a natural number, and $\epsilon \in p^{\mathbb{Z}}$. Let*

$$\mathcal{K}_{\mathfrak{p}}(\mathfrak{n}^{\mathrm{PL}}_{\geq -n}) \lhd \mathrm{Col}(X(Z_{\mathfrak{p}}))$$

*denote the ideal which defines the Chabauty–Kim locus $X(Z_{\mathfrak{p}})_n$; we refer to $\mathcal{K}_{\mathfrak{p}}(\mathfrak{n}^{\mathrm{PL}}_{\geq -n})$ as the* $p$-adic Chabauty–Kim ideal associated to $\mathfrak{n}^{\mathrm{PL}}_{\geq -n}$:

(1) *Suppose* $\mathcal{A}_{\mathrm{Loci}}(Z, p, n, \epsilon)$ *halts. Then there are functions* $\{F_i^{\mathfrak{p}}\}$ *generating the* $\mathfrak{p}$*-adic Chabauty–Kim ideal* $\mathcal{K}_{\mathfrak{p}}(\mathfrak{n}^{\mathrm{PL}}_{\geq -n})$ *associated to* $\mathfrak{n}^{\mathrm{PL}}_{\geq -n}$*, such that*

$$|\tilde{F}_i^{\mathfrak{p}} - F_i^{\mathfrak{p}}| < \epsilon$$

*for all* $i$*.*

(2) *Suppose Zagier's conjecture* (Conjecture 2.2.5) *holds for* $K$ *and* $n' \leq n$*. Suppose Goncharov exhaustion* (Conjecture 2.2.7) *holds for* $Z$ *and* $n' \leq n$*. Suppose the period conjecture holds for the open subscheme* $Z^o \subset Z$ *constructed in segment 3.8 in half-weights* $n' \leq n$*. Suppose* $K$ *obeys the Hasse principle for finite cohomology* (Condition 2.2.13) *in half-weights* $2 \leq n' \leq n$*. Then the computation* $\mathcal{A}_{\mathrm{Loci}}(Z, p, n, \epsilon)$ *halts.*

We remark that part (1) of the theorem is independent of the choice of norm on the space of polylogarithmic functions up to an admissible change in $\epsilon$. We complete the construction of the Loci algorithm and prove Theorem 2.4.1 in Section 4.

## 3. Construction of arithmetic algorithms

### 3.1. *Generators for graded free algebras.*

**Proposition 3.1.1.** *Let* $S = \bigcup_{i=1}^{\infty} S_i$ *be a disjoint union of finite sets, and similarly*

$$S' = \bigcup_{i=1}^{\infty} S'_i.$$

*Let* $k$ *be a field and* $k[S], k[S']$ *associated graded free algebras and* $I, I'$ *the augmentation ideals. Let*

$$\phi : k[S'] \to k[S]$$

*be a homomorphism which preserves the grading. Suppose the induced map*

$$I'/I'^2 \to I/I^2$$

*is iso. Then* $\phi$ *is iso.*

*Proof.* For $n \geq 1$, we have $I_n = k[S]_n$. Surjectivity follows by induction using the short exact sequences

$$0 \to (I^2)_n \to k[S]_n \to (I/I^2)_n \to 0.$$

Since $S_i$ maps to a basis of $(I/I^2)_i$, the bijection

$$(I'/I'^2)_i \to (I/I^2)_i$$

gives us a bijection between $S'_i$ and $S_i$. For any $n$, $\phi$ maps $S'_{\leq n}$ into $k[S]_{\leq n}$, so $\phi$ restricts to a map

$$k[S'_{\leq n}] \to k[S_{\leq n}]$$

of subalgebras generated in degrees $\leq n$. These are surjective maps of polynomial algebras of same finite Krull dimension. This means that

$$\operatorname{Spec} k[S_{\leq n}] \to \operatorname{Spec} k[S'_{\leq n}]$$

is a closed immersion between affine spaces of same dimension, hence an isomorphism by the Hauptidealsatz. □

### 3.2. *Generators for mixed Tate groups.*

**3.2.1.** For a review of free prounipotent groups, we refer the reader to Section 2 of [Dan-Cohen and Wewers 2016]. By a *mixed Tate group* over a field $k$ of characteristic zero, we mean a free prounipotent group $U$ equipped with a grading of the Lie algebra

$$\mathfrak{n} = \operatorname{Lie} U$$

such that $\mathfrak{n}_i = 0$ for $i \geq 0$. The Lie algebra $\mathfrak{n}$ admits a set of homogeneous free generators, and we define a *set of homogeneous free generators of* $U$ to be a set of homogeneous free generators of $\mathfrak{n}$. The grading on $\mathfrak{n}$ induces also a grading of the completed universal enveloping algebra $\mathcal{U} = \mathcal{U}\mathfrak{n}$ such that $\mathcal{U}_0 = k$ and $\mathcal{U}_i = 0$ for $i > 0$, as well as a grading on the coordinate ring $A = \mathcal{O}(U) = \mathcal{U}^\vee$ such that $A_0 = k$ and $A_i = 0$ for $i < 0$. We refer to the graded degree of an element (of $\mathfrak{n}$, $\mathcal{U}$, $A$) as its *half-weight*.

The kernel of the comultiplication

$$E_n < A_n$$

is the space of *extensions*. Indeed, by the general theory of mixed Tate categories we have exact sequences

$$0 \to \operatorname{Ext}^1_{\operatorname{Rep}(\mathbb{G}_m \ltimes U)}(k(0), k(n)) \to A_n \to \bigoplus_{\substack{i+j=n \\ i,j\geq 1}} A_i \otimes A_j$$

where $k(i)$ denotes the trivial $U$-representation in half-weight $-i$. Similarly, the multiplication gives rise to a subspace

$$A_n > D_n,$$

namely the image of the map

$$A_n \leftarrow \bigoplus_{\substack{i+j=n \\ i,j\geq 1}} A_i \otimes A_j;$$

we refer to $D_n$ as the *space of decomposable elements*.

**Proposition 3.2.2.** *Let $U$ be a mixed Tate group, and let $A$ denote its coordinate ring. For each $i$ let $E_i$ denote the space of extensions in $A_i$, $D_i$ the space of decomposable elements. Let $\mathcal{P}_i$ be a linearly independent subset of $A_i$ which spans a subspace $P_i$ complementary to $E_i + D_i$. Let $\mathcal{E}_i$ be a basis for $E_i$ and let $\mathcal{E} = \bigcup \mathcal{E}_i$, $\mathcal{P} = \bigcup \mathcal{P}_i$. Then as a ring,*

$$A = k[\mathcal{E} \cup \mathcal{P}].$$

*Proof.* The subspaces $E_i$ and $D_i$ are disjoint. To see this, fix an arbitrary set

$$\epsilon' = \bigcup_{i=1}^{\infty} \epsilon'_{-i}$$

of homogeneous free generators for $U$, and for $w$ a word in $\epsilon'$, let $f_w \in A$ denote the associated function. Then $E_i$ has basis

$$\{f_a \mid a \in \epsilon'_{-i}\}$$

dual to the set of one-letter words of half-weight $-i$, while $D_i$ is spanned by shuffle products of functions $f_w$ with $w$ a word in $\epsilon'_{>-i}$, so is contained in the space with basis

$$\{f_w \mid w \in \mathrm{Words}_{-i}(\epsilon'_{>-i})\}.$$

It follows that $A_i$ decomposes as a direct sum

$$A_i = E_i \oplus P_i \oplus D_i \tag{3.2.2$*$}$$

and that $\mathcal{E}_i \cup \mathcal{P}_i$ maps to a basis of $(I/I^2)_i$. Hence, by Proposition 3.1.1, $\mathcal{E} \cup \mathcal{P}$ forms a set of free $k$-algebra generators for $A$. □

**Proposition 3.2.3.** *In the situation and the notation of Proposition 3.2.2, let*

$$\epsilon_{-i} \subset \mathcal{U}_{-i}$$

*be the set of elements dual to the elements of $\mathcal{E}_i$ relative to the decomposition* (3.2.2$*$)*. Then*

$$\epsilon := \bigcup_{i=1}^{\infty} \epsilon_{-i}$$

*forms a set of free generators for $U$.*

*Proof.* We claim that every element

$$\epsilon_{-i,j} \in \epsilon_{-i}$$

is of Lie type; if

$$\nu : \mathcal{U} \to \mathcal{U} \otimes \mathcal{U}$$

denotes the comultiplication, then

$$\nu(\epsilon_{-i,j}) = 1 \otimes \epsilon_{-i,j} + \epsilon_{-i,j} \otimes 1. \tag{$*$}$$

Let

$$\mathcal{P}'_i = \mathcal{E}_i \cup \mathcal{P}_i.$$

According to Proposition 3.2.2, the set $\mathcal{D}_i$ of monomials in $\mathcal{P}'_{<i}$ forms a basis of $D_i$. Let

$$\mathcal{A}_i = \mathcal{E}_i \cup \mathcal{P}_i \cup \mathcal{D}_i.$$

It suffices to check the equality $(*)$ after pairing with an arbitrary basis element

$$\mathcal{A}_{i',j'} \otimes \mathcal{A}_{i'',j''}$$

of $A_{i'} \otimes A_{i''}$ with $i' + i'' = i$. We have

$$\begin{aligned}\langle \nu(\epsilon_{-i,j}), \mathcal{A}_{i',j'} \otimes \mathcal{A}_{i'',j''} \rangle &= \langle \epsilon_{-i,j}, \mathcal{A}_{i',j'} \cdot \mathcal{A}_{i'',j''} \rangle \\ &= \begin{cases} 1 & \text{if } \mathcal{A}_{i',j'} = 1 \text{ and } \mathcal{A}_{i'',j''} = \mathcal{E}_{i,j} \text{ is dual to } \epsilon_{-i,j}, \\ 1 & \text{if } \mathcal{A}_{i',j'} = \mathcal{E}_{i,j} \text{ and } \mathcal{A}_{i'',j''} = 1, \\ 0 & \text{otherwise.} \end{cases} \\ &= \langle 1 \otimes \epsilon_{-i,j}, \mathcal{A}_{i',j'} \otimes \mathcal{A}_{i'',j''} \rangle + \langle \epsilon_{-i,j} \otimes 1, \mathcal{A}_{i',j'} \otimes \mathcal{A}_{i'',j''} \rangle \\ &= \langle 1 \otimes \epsilon_{-i,j} + \epsilon_{-i,j} \otimes 1, \mathcal{A}_{i',j'} \otimes \mathcal{A}_{i'',j''} \rangle\end{aligned}$$

which shows that $\epsilon_{-i,j}$ is of Lie type as claimed.

It follows that $\epsilon_{-i}$ is a subset of the graded piece $\mathfrak{n}_{-i}$ of the Lie algebra $\mathfrak{n} \subset \mathcal{U}$, which maps to a basis of $\mathfrak{n}^{\mathrm{ab}}_{-i}$. It follows that $\epsilon$ forms a set of free generators as stated. □

**3.3.** Recall that by an *open integer scheme* we mean on open subscheme

$$Z \subset \operatorname{Spec} \mathcal{O}_K,$$

$K$ a number field. By a *number scheme* we mean $\operatorname{Spec} K$, $K$ a number field. Given $Z$ an open integer or number scheme, we let

$$A(Z) = \mathcal{O}(U(Z))$$

denote the graded Hopf algebra of unramified mixed Tate motives over $Z$.

**3.4.** Given an open integer scheme $Z$ with function field $K$ and a unipotent iterated integral $I^b_a(c_1, \dots, c_r) \in A(K)_n$, we say that $I$ is *combinatorially unramified over* $Z$ if the associated reduced divisor

$$D = \{a, b, c_1, \dots, c_r\}$$

is étale over $Z$. We denote the $\mathbb{Q}$-vector space of formal linear combinations of such tuples $(a; c_1, \dots, c_r; b)$ by $\mathrm{CUI}(Z)_r$, the space of *formal integrands in half-weight* $r$.

**3.5.** If $k$ is a field equipped with an absolute value $|\cdot|$, we say that a subset of $k^n$ is $\epsilon$-linearly independent if each of the associated determinants has absolute value greater than $\epsilon$.

**3.6.** Let $Z$ be an open integer scheme and $p \in \mathbb{Z}$ a prime such that $Z$ is totally split above $p$. Recall that $A(\mathcal{O}_\mathfrak{p})$ denotes the graded Hopf algebra of mixed Tate filtered $\phi$ modules over $K_\mathfrak{p}$, and recall that $A(\mathcal{O}_\mathfrak{p})$ possesses a *standard basis*. We say that a subset

$$\mathcal{P} \subset A(Z)_n$$

is *$\epsilon$-linearly independent relative to* $\mathfrak{R}_p$ if its image in $\prod_{\mathfrak{p} \mid p} A(\mathcal{O}_\mathfrak{p})_n$ is *$\epsilon$-linearly independent* with respect to the standard basis.

### 3.7. *Realization algorithm.*

**3.7.1.** We recall from segment 2.2.9 that $U(\mathbb{Z}_p)$ denotes the unipotent fundamental group of the category of mixed Tate filtered $\phi$ modules, that it contains a special $\mathbb{Q}_p$-point $u$, and that the family

$$v_i = (\log u)_i \in \mathfrak{n}(\mathbb{Q}_p)$$

for $i \in \mathbb{Z}_{\leq -1}$ forms a set of free generators. The associated shuffle basis of $A(\mathbb{Z}_p)$ (which is dual to the basis of the universal enveloping algebra consisting of words in the generators) is what we call the *standard basis*. We now construct an algorithm for evaluating an iterated integral $I_a^b(\omega)$, whose associated divisor $D$ is a union of $\mathbb{Z}_p$-points, on a word

$$w = v_{-i_r} \cdots v_{-i_2} v_{-i_1}$$

in the generators $v_i$ to given precision $\epsilon$.

**3.7.2.** Let $Z \subset \operatorname{Spec} \mathcal{O}_K$ be an open integer scheme, and $\mathfrak{p} \in Z$ a prime which is totally split. Recall from segment 3.4 that $\mathrm{CUI}(Z)_r$ denotes the $\mathbb{Q}$-vector space of combinatorially unramified integrands in half-weight $r$. The *realization algorithm*, alluded to above and constructed in segment 3.7.16 below, may be interpreted as an algorithm which takes a natural number $r$ and an $\epsilon \in p^{\mathbb{Z}}$ as input, and returns a linear map

$$\widetilde{{}^{U}I^{F\phi}_{\mathrm{Stdd}}} : \mathrm{CUI}(Z)_r \to \prod_{\mathfrak{p} \mid p} A(\mathcal{O}_{\mathfrak{p}})_r$$

given explicitly by a matrix with rational entries. If

$$\mathrm{Re} : A(Z)_r \to \prod_{\mathfrak{p} \mid p} A(\mathcal{O}_{\mathfrak{p}})_r$$

denotes the realization map, and

$${}^{U}I : \mathrm{CUI}(Z)_r \to A(Z)_r$$

denotes the map taking an integrand to the associated unipotent iterated integral, then the triangle

$$\begin{array}{ccc} \mathrm{CUI}(Z)_r & & \\ \downarrow{\scriptstyle {}^{U}I} & \searrow{\scriptstyle \widetilde{{}^{U}I^{F\phi}_{\mathrm{Stdd}}}} & \\ A(Z)_r & \xrightarrow[\mathrm{Re}]{} & \prod_{\mathfrak{p} \mid p} A(\mathcal{O}_{\mathfrak{p}})_r \end{array}$$

fails to commute by at most $\epsilon$. Said differently, $\widetilde{{}^{U}I^{F\phi}_{\mathrm{Stdd}}}$ is an approximation of the matrix representing the composite

$${}^{U}I^{F\phi} := \mathrm{Re} \circ ({}^{U}I)$$

with respect to the "standard" bases on source and target. An example is worked out in segment 7.5.3 of [Dan-Cohen and Wewers 2016]. Note, however, that the triviality of the motivic Galois action on the polylogarithmic quotient makes that example deceptively simple compared to the general algorithm that follows.

We begin in segements 3.7.3–3.7.8 by deriving a formula (Lemma 3.7.4) for the action of the special $\mathbb{Q}_p$-point $u$ of the unipotent mixed Tate filtered $\phi$ Galois group $U(\mathbb{Z}_p)$ on any generator of the unipotent fundamental group of $X := \mathbb{A}^1_{\mathbb{Q}_p} \setminus D$. We then note in segment 3.7.9 that the formula of Lemma 3.7.4 gives rise to an algorithm for computing the action of $u$ on any word in the generators. In segment 3.7.11 we extend this algorithm to include not only fundamental groups but also path torsors. After a few elementary observations regarding the matrix entries of a graded representation of a graded Lie algebra on a graded vector space equipped with a graded basis on which we do not wish to impose an ordering (segements 3.7.13–3.7.14), and after constructing a certain family of polynomials with rational coefficients based on these observations (segment 3.7.15), we construct the realization algorithm in segment 3.7.16 and we state and verify the correctness of its output in segements 3.7.17–3.7.18.

**3.7.3.** Although our application is global, this algorithm may equally be constructed in a purely $p$-adic situation.[11] In order to minimize the number of decorations, we introduce notation specific to the present situation; these will remain in effect through the proof of Proposition 3.7.18.

We let $\boldsymbol{F\phi}$ denote the category of mixed Tate filtered $\phi$ modules over $\mathbb{Q}_p$. We let

$$\Omega : \boldsymbol{F\phi} \to \mathrm{Vect}(\mathbb{Q}_p)$$

denote the forgetful functor. As in segment 2.2.9, we let $U(\mathbb{Z}_p)$ denote the unipotent part of the Tannakian fundamental group $\mathrm{Aut}^{\otimes}(\Omega)$. Given $E \in \boldsymbol{F\phi}$, $v \in \rho(E)$ and $f \in \rho(E)^{\vee}$, we let

$$[E, v, f]$$

denote the function

$$U(\mathbb{Z}_p) \to \mathbb{A}^1_{\mathbb{Q}_p}$$

given on a point $\gamma$ with values in an arbitrary $\mathbb{Q}_p$-algebra by

$$\gamma \mapsto f(\gamma v).$$

Recall that $u$ denotes the $\mathbb{Q}_p$-point of $U(\mathbb{Z}_p)$ associated to the $p$-adic period map.

Let $D$ be a finite set of elements of $\mathbb{Z}_p$ no two of which are congruent modulo $p$. Let $X := \mathbb{A}^1_{\mathbb{Z}_p} \setminus D$. We consider two $\mathbb{Z}_p$-integral base points $a, b$ of $X$. We let ${}_aP_a$ denote the filtered $\phi$ realization of the unipotent fundamental group of $X_{\mathbb{Q}_p}$ at $a$, a unipotent group object of $\boldsymbol{F\phi}$. We let ${}_bP_a$ denote the filtered

[11] There's a slight caveat: where the algorithm and its subalgorithms take $p$-adic numbers an input, those must be specified by a finite amount of data. This means that the domain of the algorithm must be restricted to those $p$-adic numbers which can be specified by a finite amount of data. However, since we've agreed not to keep track of $\epsilon$, this restriction on the domain need not concern us any further.

$\phi$ realization of the unipotent path torsor. We let ${}_a\mathcal{U}_a$ denote the completed universal enveloping algebra of ${}_aP_a$ and let

$$ {}_b\mathcal{U}_a := {}_a\mathcal{U}_a \times_{{}_aP_a} {}_bP_a $$

be the associated rank one free module.

**Lemma 3.7.4.** *We put ourselves in the situation and the notation of segment 3.7.3. Consider an element $c$ of the set $D$ of punctures and a $\mathbb{Z}_p$-integral base point $a$ of $X(\mathbb{Z}_p)$. Let $e^c$ denote the element of ${}_a\mathcal{U}_a$ associated to monodromy about $c$. Then we have the equality*

$$ ue^c = p\Big(\sum_{\eta}\Big(\int_c^a \eta\Big)\eta\Big)\cdot e^c \cdot \Big(\sum_{\omega}\Big(\int_a^c \omega\Big)\omega\Big) $$

*in the noncommutative formal power series ring ${}_a\mathcal{U}_a$. Both sums run over the set of words in the family of differential forms*

$$ \left\{\frac{dt}{t-d}\right\}_{d\in D} $$

*and the integrals are regularized with respect to the unit tangent vector at $c$. The integral of the empty word is defined to be* 1.

The proof (which is purely formal) spans segements 3.7.5–3.7.8.

**3.7.5.** We recall that each path torsor ${}_bP_a$ possesses a unique $\mathbb{Q}_p$-valued point contained in step 0 of the Hodge filtration

$$ p^{\mathrm{dR}} = {}_bp_a^{\mathrm{dR}} $$

and a unique $\mathbb{Q}_p$-valued point

$$ p^{\mathrm{cris}} = {}_bp_a^{\mathrm{cris}} $$

fixed by Frobenius. We have

$$ up^{\mathrm{dR}} = p^{\mathrm{cris}}. $$

We use the de Rham path ${}_bp_a^{\mathrm{dR}}$ to identify ${}_b\mathcal{U}_a$ with ${}_a\mathcal{U}_a$. Let us write ${}_b\omega_a$ for a word $\omega$ *regarded as an element of* ${}_b\mathcal{U}_a$. Thus,

$$ {}_b\omega_a = {}_bp_a^{\mathrm{dR}} \cdot {}_a\omega_a. $$

In this notation,

$$ {}_bp_a^{\mathrm{dR}} = {}_b1_a. $$

We denote ${}_ce_c^c$ simply by $\epsilon^c$.

By Besser's definition of the $p$-adic iterated integrals

$$ \int_a^b \omega, $$

we have

$$ {}_b p_a^{\mathrm{cris}} = \sum_{\omega} \left( \int_a^b \omega \right) {}_b\omega_a $$

where the integral of the empty word is defined to be 1.

**3.7.6.** We have

$$ {}_b p_a^{\mathrm{dR}} \cdot {}_a\omega_a = {}_b\omega_b \cdot {}_b p_a^{\mathrm{dR}}. \qquad (*) $$

Indeed, this equality reduces to the case of a one-letter word

$$ \omega = e^c. $$

We have

$$ {}_b e_b^c = {}_b p_c^{\mathrm{dR}} \cdot \epsilon^c \cdot {}_c p_b^{\mathrm{dR}}. $$

Since the composition of de Rham paths is again a de Rham path, it follows that both sides of equation $(*)$ (when $\omega = e^c$) are equal to

$$ {}_b p_c^{\mathrm{dR}} \cdot \epsilon^c \cdot {}_c p_a^{\mathrm{dR}}. $$

**3.7.7.** By segment 3.7.6, we have for any word $\omega$ in the set $D$ of punctures and any points $a$, $b$, $c$

$$ {}_c\eta_b \cdot {}_b\omega_a = {}_c(\eta\omega)_a. $$

**3.7.8.** We thus have

$$ \begin{aligned} u({}_a e_a^c) &= u({}_a p_c^{\mathrm{dR}} \cdot \epsilon^c \cdot {}_c p_a^{\mathrm{dR}}) \\ &= u({}_a p_c^{\mathrm{dR}}) \cdot u(\epsilon^c) \cdot u({}_c p_a^{\mathrm{dR}}) \\ &= \left( \sum_{\eta} \left( \int_c^a \eta \right) {}_a\eta_c \right) \cdot p\epsilon^c \cdot \left( \sum_{\omega} \left( \int_a^c \omega \right) {}_c\omega_a \right) \\ &= {}_a\left( p\left( \sum_{\eta} \left( \int_c^a \eta \right) \eta \right) \cdot e^c \cdot \left( \sum_{\omega} \left( \int_a^c \omega \right) \omega \right) \right)_a. \end{aligned} $$

This completes the proof of Lemma 3.7.4.

**3.7.9.** *Action of $u$ on an arbitrary loop-word.* Let $\omega$ be a word in the set $D$ of punctures regarded as an element of the completed universal enveloping algebra ${}_a\mathcal{U}_a$ at a base-point $a$ of $X$. Since the action of $u$ on ${}_a\mathcal{U}_a$ respects multiplication, Lemma 3.7.4 coupled with the algorithm of [Dan-Cohen and Chatzistamatiou 2014] for computing $p$-adic iterated integrals provides an algorithm for computing any coefficient of the noncommutative formal power series $u\omega$ to precision $\epsilon$.

**3.7.10.** *Remark on the unipotent nature of the action on loop algebras.* We recall that the generators $e^c$ ($c \in D$) of the completed universal enveloping algebra ${}_a\mathcal{U}_a$ (or "loop algebra") have half-weight $-1$ and that the action of $U(\mathbb{Z}_p)$ (and hence of $u$) on ${}_a\mathcal{U}_a$ is unipotent with respect to the weight filtration. More specifically, the formula of Lemma 3.7.4 shows that $u\omega$ has coefficient 1 in front of the word $\omega$ itself,

and that every word which occurs with nonzero coefficient has $\omega$ as a subword (by which we mean a subsequence of not necessarily consecutive letters).

**3.7.11.** *Action of $u$ on arbitrary path-words.* Let $\omega$ be a word in the set $D$ of punctures regarded as an element of the completed universal enveloping bimodule ${}_b\mathcal{U}_a$ associated to a pair of base-points $a$ and $b$ of $X$. The algorithm of segment 3.7.9 may be upgraded to an algorithm which computes any coefficient of the noncommutative formal power series $u\omega$ to precision $\epsilon$, or, which is the same, an algorithm which computes the $p$-adic period

$$[{}_b\mathcal{U}_a, \omega, f_{\omega'}](u)$$

of any matrix entry $[{}_b\mathcal{U}_a, \omega, f_{\omega'}]$. Indeed, $[{}_b\mathcal{U}_a, \omega, f_{\omega'}](u)$ is the $\omega'$-coefficient of the noncommutative formal power series

$$u({}_b\omega_a) = u({}_b p_a^{\mathrm{dR}} \cdot {}_a\omega_a) = {}_b p_a^{\mathrm{cris}} \cdot u({}_a\omega_a). \qquad (*)$$

Using the algorithm of segment 3.7.9, we expand $u({}_a\omega_a)$ as a linear combination of words

$$u({}_a\omega_a) = \sum_\eta c_\eta \cdot {}_a\eta_a$$

with coefficients $c_\eta \in \mathbb{Q}_p$ computed to $p$-adic precision $\epsilon$. We then have

$$\begin{aligned}
[{}_b\mathcal{U}_a, \omega, f_{\omega'}](u) &= f_{\omega'}(u({}_b\omega_a)) \\
&= f_{\omega'}({}_b p_a^{\mathrm{cris}} \cdot u({}_a\omega_a)) \\
&= f_{\omega'}({}_b p_a^{\mathrm{cris}} \cdot \sum_\eta c_\eta \cdot {}_a\eta_a) \\
&= \sum_\eta c_\eta f_{\omega'}({}_b p_a^{\mathrm{cris}} \cdot {}_a\eta_a) \\
&= \sum_\eta c_\eta \int_a^b \omega'/\eta
\end{aligned}$$

where the right-division $\omega'/\eta$ is defined to be zero whenever $\omega'$ is not right-divisible by $\eta$. Using [Dan-Cohen and Chatzistamatiou 2014] again we compute these last $p$-adic iterated integrals to precision $\epsilon$.

**3.7.12.** *Remark on the unipotent nature of the action on path modules.* We recall that the generators $e^c$ ($c \in D$) of the completed universal enveloping bimodule ${}_b\mathcal{U}_a$ (or "path module") have half-weight $-1$ and that the action of $U(\mathbb{Z}_p)$ (and hence of $u$) on ${}_b\mathcal{U}_a$ is unipotent with respect to the weight filtration. This squares with the computation of segment 3.7.11. To see this more clearly, we repeat the computation in slightly different notation: we have

$$u({}_b\omega_a) = {}_b p_a^{\mathrm{cris}} \cdot u({}_a\omega_a) = \sum_\theta \left( \int_a^b \theta \right) {}_b\theta_a \cdot \sum_\eta c_{\eta a} \eta_a = \sum_{\theta,\eta} \left( c_\eta \int_a^b \theta \right) {}_b\theta\eta_a.$$

We find that the coefficient in front of the word $\omega = \eta$ ($\theta = 1$) is 1, and that all words occurring in the sum are left-multiples of words which contain $\omega$ as a subword.

**3.7.13.** *Elementary remarks on matrices with respect to unordered bases.* Let $k$ be a field, $V$, $W$ finite dimensional vector spaces,

$$\phi : V \to W$$

a linear map, $\mathcal{V}$ a basis of $V$ and $\mathcal{W}$ a basis of $W$. Then the associated matrix is indexed by the set $\mathcal{V} \times \mathcal{W}$. The entry associated to the pair $(v, w)$ is given by

$${}_w\phi_v = w^\vee(\phi v)$$

where $w^\vee$ is the linear functional on $W$ dual to $w$ with respect to the basis $\mathcal{W}$.

If $V = \bigoplus_i V_i$, $W = \bigoplus_j W_j$ are finite direct sums of finite dimensional vector spaces with bases $\mathcal{V}_i$, $\mathcal{W}_j$ and $\phi = \bigoplus_{i,j} \phi_{i,j}$ is a direct sum of linear maps

$$\phi_{i,j} : V_i \to W_j,$$

then the matrix associated to $\phi$ is given in terms of the matrices of the $\phi_{i,j}$ as follows: if $v \in \mathcal{V}_i$ and $w \in \mathcal{W}_j$ then

$${}_w\phi_v = {}_w(\phi_{i,j})_v.$$

**3.7.14.** *Elementary remarks on graded pieces of graded representations.* Let $\mathfrak{g} = \bigoplus_n \mathfrak{g}_n$ be a graded Lie algebra over a field $k$, $E = \bigoplus E_i$ a finite dimensional graded vector space,

$$\rho : \mathfrak{g} \to \mathfrak{gl}\, E$$

a graded representation. Let $\mathcal{U}$ denote the universal enveloping algebra of $\mathfrak{g}$. Then the induced ring homomorphism

$$\rho : \mathcal{U} \to \operatorname{End} E$$

preserves gradings. We spell out what this means. We let

$$\operatorname{End}^n E \subset \operatorname{End} E$$

denote the subspace of homomorphisms which are graded of graded degree $n$:

$$\operatorname{End}^n E = \bigoplus_i \operatorname{Hom}(E_i, E_{i+n}).$$

Then $\phi$ sends the $n$-th graded piece $\mathcal{U}_n$ of $\mathcal{U}$ into $\operatorname{End}^n E$. This also means that $\rho$ is compatible with projections, in the sense that the squares

$$\begin{array}{ccc} \mathcal{U} & \longrightarrow & \operatorname{End} E \\ \downarrow & & \downarrow \\ \mathcal{U}_n & \longrightarrow & \operatorname{End}^n E \end{array}$$

commute.

In terms of matrices, the projection has the effect of setting entries in all other graded degrees equal to zero. More precisely, if $\phi \in \mathrm{End}\, E$ has $n$-th graded piece $\phi^n \in \mathrm{End}^n E$, if $\mathcal{E} = \bigcup_i \mathcal{E}_i$ is a graded basis of $E$ and if $v \in \mathcal{E}_i$, $w \in \mathcal{E}_j$ are basis vectors of graded degrees $i$ and $j$, respectively, then

$$ {}_w\phi_v^n = \begin{cases} {}_w\phi_v & \text{if } j - i = n, \\ 0 & \text{otherwise.} \end{cases} $$

To see this, let $P_i$ denote the idempotent

$$ E \twoheadrightarrow E_i \hookrightarrow E $$

associated to the $i$-th graded piece. Then

$$ \phi^n(v) = P_{i+n}\phi(v), $$

so

$$ {}_w\phi_v^n = w^\vee P_{i+n}\phi(v). $$

We complete the verification by noting that

$$ w^\vee \circ P_{i+n} = \begin{cases} w^\vee & \text{if } j = i + n, \\ 0 & \text{otherwise.} \end{cases} $$

**3.7.15.** *Universal polynomials for entries of products of graded pieces of the logarithm of a matrix.* Let $\mathrm{Word}(D)$ denote the set of words in the set $D$ of punctures. Let $R$ be the polynomial $\mathbb{Q}$-algebra

$$ R = \mathbb{Q}[\mathrm{Word}\, D \times \mathrm{Word}\, D] $$

graded by setting the degree of a word equal to minus its length as usual. We denote the generator associated to a pair of words $\omega$, $\eta$ by $x_{\omega,\eta}$. Let $M$ be the Word $D \times$ Word $D$-matrix whose $(\omega, \eta)$-th entry ${}_\eta M_\omega$ is 1 if $\omega = \eta$, $x_{\omega,\eta}$ if $\eta$ contains $\omega$ as a subword, and 0 otherwise. Then the logarithm of $M$ converges (in the sense that each entry is a polynomial). We let

$$ N = \log M. $$

For any integer $l$, we define a new matrix ${}^l N$ with entries

$$ {}^l_\eta N_\omega = \begin{cases} {}_\eta N_\omega & \text{if } |\eta| - |\omega| = l, \\ 0 & \text{otherwise.} \end{cases} $$

Let $\omega$ be a word of length $n$ in $D$ and let

$$ w = l_1 \cdots l_r $$

be a word in the set $\mathbb{Z}_{<0}$ of negative integers such that

$$ l_1 + \cdots + l_s = -n. $$

We define a polynomial $P(\omega, w) \in R$ by taking the $(\varnothing, \omega)$-th entry

$$P(\omega, w) = {}_{\omega}[{}^{l_1}N \cdots {}^{l_s}N]_1$$

of the product of graded pieces of $N$ indicated by $\omega$.

**3.7.16.** *Main algorithm.* We now arrive at the construction of the realization algorithm. As input, the algorithm takes a positive real number $\epsilon$, a prime number $p$, a finite set $D$ of elements of $\mathbb{Z}_p$ and two further elements $a, b$ (given up to $p$-adic precision $\epsilon$), a natural number $n$, a word $\omega$ of length $n$ in the set $D$, and a word $w$ of degree $n$ in the set of symbols

$$\{v_{-1}, v_{-2}, v_{-3}, \ldots\}$$

indexed and weighted by the negative integers. Distinct points in the set $D \cup \{a, b\}$ must not be congruent modulo $p$, but the elements $a, b$ may belong to the set $D$ (in the latter case, they will be treated as unit tangent vectors). The output consists of a single element

$$\mathcal{A}_{\mathrm{Real}}(\epsilon, p, D, a, b, \omega, w)$$

of $\mathbb{Q}_p$ given up to $p$-adic precision $\epsilon$. To construct it, we first check which variables $x_{\theta,\eta}$ intervene in the polynomial $P(\omega, w)$ constructed in segment 3.7.15. For each such variable, we apply the subalgorithm of segment 3.7.11 to compute the element

$$[{}_b\mathcal{U}_a, \eta, f_\theta](u)$$

of $\mathbb{Q}_p$ to precision $\epsilon$. We then output the value

$$\mathcal{A}_{\mathrm{Real}}(\epsilon, p, D, a, b, \omega, w) = P(\omega, w)(\{[{}_b\mathcal{U}_a, \eta, f_\theta](u)\}_{\eta,\theta}).$$

**3.7.17.** We now announce the meaning of the output. In terms of the input, we let $X = \mathbb{A}^1_{\mathbb{Q}_p} \setminus D$, and we work with the filtered $\phi$ unipotent path bimodule ${}_b\mathcal{U}_a$ on $X$. Let $n$ be the length of the word $\omega$. As in segment 4.9 of [Dan-Cohen and Wewers 2016], we define the *unipotent filtered $\phi$ iterated integral* $I^b_a(\omega) \in A(\mathbb{Z}_p)_n$ to be the Tannakian matrix entry

$$I^b_a(\omega) = [{}_b\mathcal{U}_a, {}_b1_a, f_\omega].$$

Via the isomorphism

$$A(\mathbb{Z}_p)_n = \mathcal{U}(\mathbb{Z}_p)^\vee_{-n},$$

the unipotent filtered $\phi$ iterated integral $I^b_a(\omega)$ may be *evaluated* at the element $w \in \mathcal{U}(\mathbb{Z}_p)_{-n}$.

**Proposition 3.7.18.** *The* realization algorithm $\mathcal{A}_{\mathrm{Real}}$ *halts. Moreover, in the notation of segment 3.7.17, its output is within $\epsilon$ of the $p$-adic number $I^b_a(\omega)(w)$.*

*Proof.* The halting presents no issue. Turning to the verification of the correctness of the output, we fix an arbitrary input datum

$$(\epsilon, p, D, a, b, \omega, w)$$

and we set ourselves the task of computing $I_a^b(\omega)(w)$ in terms of the periods

$$[{}_b\mathcal{U}_a, \eta, f_\theta](u);$$

this is mostly a matter of rearranging definitions. Let

$${}_b\mathcal{U}_a^{\geq -n} = {}_b\mathcal{U}_a / {}_b\mathcal{U}_a I^{n+1}$$

where $I$ denotes the augmentation ideal of ${}_a\mathcal{U}_a$. The quotient module ${}_b\mathcal{U}_a^{\geq -n}$ has vector space basis consisting of words of length $\leq n$. We note that

$$[{}_b\mathcal{U}_a^{\leq -n}, \eta, f_\theta](u) = [{}_b\mathcal{U}_a, \eta, f_\theta](u)$$

so long as $\eta$ and $\theta$ are both of length $\leq n$.

Let ${}_b\rho_a$ denote homomorphism of graded $\mathbb{Q}_p$-algebras

$$\mathcal{U}(\mathbb{Z}_p) \to \mathrm{End}\, {}_b\mathcal{U}_a^{\leq -n}$$

induced by the action of the filtered $\phi$ Galois group $U(\mathbb{Z}_p)$ on the path bimodule ${}_b\mathcal{U}_a$. Then

$$I_a^b(\omega)(w) = {}_\omega[{}_b\rho_a(w)]_1$$

is the $(1, \omega)$-th entry of the matrix associated to the endomorphism ${}_b\rho_a(w)$ of ${}_b\mathcal{U}_a^{\leq -n}$. To compute it, write $w$ as a product of letters

$$w = l_1 \cdots l_s$$

which we identify with the negative integers which parametrize them (while taking care *not* to confuse the above juxtaposition of letters with the product of integers). Meanwhile, recall that for $i$ a negative integer,

$$v_i = {}^i(\log u)$$

is the $i$-th graded piece of $\log u$. Thus, if we set $M$ equal to the matrix associated to ${}_b\rho_a(u)$, and $N = \log M$, we have

$${}_b\rho_a(v_i) = {}^i N.$$

Consequently,

$${}_b\rho_a(w) = {}_b\rho_a(l_1) \cdots {}_b\rho_a(l_s) = {}^{l_1}N \cdots {}^{l_s}N.$$

Putting the pieces back together, we have

$$I_a^b(\omega)(w) = {}_\omega[{}_b\rho_a(w)]_1 = P(\omega, w)(\{{}_\theta M_\eta\}_{\eta,\theta}) \sim_\epsilon P(\omega, w)(\{{}_\theta \widetilde{M}_\eta\}_{\eta,\theta}) = \mathcal{A}_{\mathrm{Real}}(\epsilon, p, D, a, b, \omega, w)$$

where ${}_\theta\widetilde{M}_\eta$ denotes the $\epsilon$-approximation produced by the algorithm of segment 3.7.11 and $\sim_\epsilon$ signals an error bounded by $\epsilon$ (up to an admissible change in $\epsilon$). □

### 3.8. *Basis algorithm.*

**3.8.1.** We now construct an algorithm which takes as input an open integer scheme

$$Z \subset \operatorname{Spec} \mathcal{O}_K,$$

a prime $p$ of $\mathbb{Z}$, a natural number $n$, and an

$$\epsilon \in p^{\mathbb{Z}},$$

and returns the following data:

(1) An open subscheme $Z^o \subset Z$. We write

$$\bar{S} = \{\mathfrak{q}_1, \dots, \mathfrak{q}_{\bar{s}}\}$$

for its complement.

(2) Sets

$$\mathcal{E}_1^g = \{\log^U \alpha_{1,1}, \dots, \log^U \alpha_{1,r_1+r_2-1}\}, \quad \mathcal{E}_1^r = \{\log^U \beta_1, \dots, \log^U \beta_{\bar{s}}\}$$

of unipotent logarithms of elements of $\mathcal{O}^*_{Z^o}$.

(3) For each integer $n' \in [2, n]$,

(a) a set of single-valued unipotent polylogarithms

$$\tilde{\mathcal{E}}_{n'} = \{\mathrm{Li}^{U,sv}_{n'}(a_{n',1}), \dots, \mathrm{Li}^{U,sv}_{n'}(a_{n',e_{n'}})\},$$

where $e_m$ denotes the dimension of the motivic extension space

$$\operatorname{Ext}^1_{Z^o}(\mathbb{Q}(0), \mathbb{Q}(m)),$$

(b) a set $\mathcal{P}_{n'}$ of unipotent iterated integrals of half-weight $n'$,

(c) an $\epsilon' \in p^{\mathbb{Z}}$,

(d) an algorithm which takes a pair $I, J$ of unipotent iterated integrals of half-weight $n'$ as input and returns a rational number

$$\langle I, J \rangle_{\epsilon'} \in \mathbb{Q}.$$

We denote this algorithm by $A_{\mathrm{Basis}}$. We first announce the meaning of its output in Proposition 3.8.2; we then construct the algorithm in segements 3.8.3–3.8.11, and prove the proposition in segements 3.8.12–3.8.14.

**Proposition 3.8.2.** (1) *Suppose $A_{\mathrm{Basis}}(Z, p, n, \epsilon)$ halts. Then we have*:

(a) *$\mathcal{E}_1^g$ forms a basis of $A(\operatorname{Spec} \mathcal{O}_K)_1$.*

(b) *$\mathcal{E}_1^g \cup \mathcal{E}_1^r$ forms a basis of $A(Z^o)_1$.*

(c) *Each $\tilde{\mathcal{B}}_{n'} := \tilde{\mathcal{E}}_{n'} \cup \mathcal{P}_{n'}$ $(n' = 2, 3, \dots, n)$ forms a basis for a subspace $\tilde{B}_{n'}$ of $A(Z^o)_{n'}$ complementary to the space $D_{n'}$ of decomposables. Moreover, the space $P_{n'}$ spanned by $\mathcal{P}_{n'}$ is disjoint from the space $E_{n'}$ of extensions.*

(d) *Relative to this basis, the projection* $\mathcal{E}_{n'}$ *of* $\tilde{\mathcal{E}}_{n'}$ *onto* $E_{n'}$ *forms a basis of* $E_{n'}$.

(e) *We let* $\mathcal{B}_{n'} = \mathcal{E}_{n'} \cup \mathcal{P}_{n'}$, *we let* $\mathcal{D}_{n'}$ *denote the set of monomials in* $\mathcal{B}_{<n'}$, *we let*

$$\mathcal{A}_{n'} = \mathcal{B}_{n'} \cup \mathcal{D}_{n'},$$

*and we denote by* $|\cdot|_{\mathcal{A}}$ *the norm induced on* $A(Z^o)_{n'}$ *by the basis* $\mathcal{A}_{n'}$. *If* $\mathrm{Li}_{n'}^{E,sv}(a_{n',i})$ *denotes the projection of* $\mathrm{Li}_{n'}^{U,sv}(a_{n',i})$ *onto* $E_{n'}$ *then we have*

$$|\mathrm{Li}_{n'}^{U,sv}(a_{n',i}) - \mathrm{Li}_{n'}^{E,sv}(a_{n',i})|_{\mathcal{A}} < \epsilon'.$$

(f) *We have* $\epsilon' \leq \epsilon$.

(g) *If* $I$, $J$ *are unipotent iterated integrals of half-weight* $n'$, *and*

$$\langle I, J \rangle_{\mathcal{A}}$$

*denotes the inner product in which the basis* $\mathcal{A}_{n'}$ *is orthonormal, then*

$$|\langle I, J \rangle_{\epsilon'} - \langle I, J \rangle_{\mathcal{A}}|_p < \epsilon'.$$

*In other words, the algorithm produced as part* (*d*) *of the output of* $A_{\mathrm{Basis}}$ *computes this inner product up to precision* $\epsilon'$.

(2) *If Zagier's conjecture* (Conjecture 2.2.5), *Goncharov exhaustion* (Conjecture 2.2.7) *and the Hasse principle for finite cohomology* (Condition 2.2.13) *hold for* $n$, $Z$, *and* $K$, *then the computation* $A_{\mathrm{Basis}}(Z, p, n, \epsilon)$ *halts.*

**3.8.3.** We write $d_G$ for the reduced Goncharov coproduct, regarded as a map

$$\mathrm{CUI}(Z)_r \to (\mathrm{CUI}(Z)_{>0})_r^{\otimes 2}.$$

**3.8.4.** The algorithm $A_{\mathrm{Basis}}$ searches arbitrarily through the countably infinite set of data $(Z^o, \tilde{\mathcal{E}}_{\leq n}, \mathcal{P}_{\leq n}, \epsilon')$. For the rest of the construction, we fix such a datum, and construct an algorithm which returns a boolean argument, as well as a function $\langle \cdot, \cdot \rangle_{\epsilon'}$. If the boolean result if *False*, we start over with a new datum. If the boolean result is *True*, we output

$$(Z^o, \tilde{\mathcal{E}}_{\leq n}, \mathcal{P}_{\leq n}, \epsilon', \langle \cdot, \cdot \rangle_{\epsilon'}).$$

For the base case with $n' = 1$ we require our basis to be of the form given in the proposition, with

$$\{a_{1,1}, \ldots, a_{1,r_1+r_2-1}\}$$

a basis for $\mathcal{O}_K^*$, and each $b_i$ a generator for a power of $\mathfrak{q}_i$.

**3.8.5.** We assume for a recursive construction that conditions (a)–(f) have been verified in half-weights $< n'$, and that the algorithm computing the inner products $\langle I, J \rangle_{\epsilon'}$ has been constructed in half-weights $< n'$. The inner products give us maps

$$\widetilde{{}^{U}I_{\tilde{A}}} : \mathrm{CUI}(Z^o)_r \to A(Z^o)_r \quad \text{and} \quad \widetilde{{}^{U}I_{\tilde{A}}}^{\otimes 2} : \mathrm{CUI}(Z^o)_r^{\otimes 2} \to A(Z^o)_r^{\otimes 2}$$

in the form of explicit matrices with rational coefficients with respect to the bases $\mathcal{A}_{<n'}$ of iterated integrals already constructed in lower half-weights.

**3.8.6.** We check if the divisors associated to the iterated integrals in $\mathcal{B}_{n'}$ are étale over $Z^o$; if not, we return *false*.

**3.8.7.** We check each element $I$ of $\tilde{\mathcal{E}}_{n'}$ for proximity to $E_{n'}$. To do so, we lift $I$ to an integrand $w \in \mathrm{CUI}(Z^o)_{n'}$, compute $\widetilde{{}^{U}I_{\tilde{\mathcal{A}}}}(d_G(w))$, and check that the $p$-adic norm of the resulting vector is $< \epsilon'$. If not, we return *False*.

**3.8.8.** We check

$$\widetilde{{}^{U}I^{F\phi}_{\mathrm{Stdd}}}(\tilde{\mathcal{E}}_{n'})$$

for $\epsilon'$-linear independence in the sense of segements 3.5 and 3.6 using the *realization algorithm* of segment 3.7. If this fails, we return *False*.

**3.8.9.** We check $d(\mathcal{P}_{n'} \cup \tilde{\mathcal{D}}_{n'})$ for $\epsilon'$-linear independence by lifting $\mathcal{P}_{n'}$, $\tilde{\mathcal{D}}_{n'}$ to $\mathrm{CUI}(Z^o)$ and applying $\widetilde{{}^{U}I_{\tilde{\mathcal{A}}}}^{\otimes 2} \circ d_G$. If this fails, we return *False*.

**3.8.10.** We check that $\epsilon'$ is sufficiently small compared to the spread of the basis $\tilde{\mathcal{A}}_{n'}$ that the projection onto the space $E_{n'}$ of extensions will preserve the linear independence of $\tilde{\mathcal{A}}_{n'}$. If this fails, we return *False*. Otherwise we return *True*.

**3.8.11.** For the inner product in half-weight $n'$, it suffices to construct

$$\langle w, x \rangle_{\epsilon'}$$

for $x \in \mathrm{CUI}(Z^o)_{n'}$ arbitrary and $w \in \tilde{\mathcal{A}}_{n'}$ a basis element. We first construct the inner products

$$\{\langle w, x \rangle \mid w \in \mathcal{P}_{n'} \cup \tilde{\mathcal{D}}_{n'}\}.$$

It may happen that $\widetilde{{}^{U}I_{\tilde{\mathcal{A}}}}^{\otimes 2}(d_G(w))$ is not in the span $V$ of the set

$$\mathcal{V} := \widetilde{{}^{U}I_{\tilde{\mathcal{A}}}}^{\otimes 2}(d_G(\mathcal{P}_{n'} \cup \tilde{\mathcal{D}}_{n'})). \tag{$*$}$$

Nevertheless, we may compute the projection $w'$ of $w$ onto $V$ with respect to the basis $(\tilde{\mathcal{A}}^{\otimes_2})_{n'}$. Subsequently, expanding $w'$ in the set $\mathcal{V}$ is a matter of solving a system of linear equations with rational coefficients; the linear independence of $(*)$ established in step 3.8.9 above, implies the uniqueness of the solution.

To compute the remaining inner products

$$\{\langle w, x \rangle_{\epsilon'} \mid w \in \tilde{\mathcal{E}}_{n'}\},$$

we replace $x$ by

$$x' = x - \sum_{w \in \mathcal{P}_{n'} \cup \tilde{\mathcal{D}}_{n'}} \langle x, w \rangle_{\epsilon'} w.$$

We then compute the projection of $\widetilde{{}^{U}I^{F\phi}_{\mathrm{Stdd}}}(x')$ onto the span of $\widetilde{{}^{U}I^{F\phi}_{\mathrm{Stdd}}}(\tilde{\mathcal{E}}_{n'})$ inside $\prod_{\mathfrak{p}\,|\,p} A(\mathcal{O}_{\mathfrak{p}})$. The linear independence of the latter, established in segment 3.8.8 above, ensures that the resulting system of linear equations will have a unique solution.

This completes the construction of the algorithm.

**3.8.12.** We now prove Proposition 3.8.2. Suppose as in part (1) of the proposition that $A_{\mathrm{Basis}}(Z, p, n, \epsilon)$ halts. Parts (a) and (b) are clear. For parts (c) and (d) we note that if $\epsilon$-approximations are $\epsilon$-linearly independent, then the actual vectors are linearly independent. Part (e) is clear, except for perhaps the admissibility of the change in $\epsilon'$; see segment 3.8.13 below. For part (f) we of course limit ourselves to searching through data satisfying $\epsilon' \leq \epsilon$ in the first place.

For part (g), we note that the square below, left, commutes.

$$\begin{array}{ccc} A(Z^o)_r & \xrightarrow{d} & (A(Z^o)^{\otimes 2})_r \\ {\scriptstyle {}^{U}I}\uparrow & & \uparrow{\scriptstyle {}^{U}I^{\otimes 2}} \\ \mathrm{CUI}(Z^o)_r & \xrightarrow[d_G]{} & (\mathrm{CUI}(Z^o)^{\otimes 2})_r \end{array} \qquad \begin{array}{ccc} A(Z^o)_r & \xrightarrow{d} & (A(Z^o)^{\otimes 2})_r \\ {\scriptstyle \widetilde{{}^{U}I_{\tilde{\mathcal{A}}}}}\uparrow & & \uparrow{\scriptstyle \widetilde{{}^{U}I_{\tilde{\mathcal{A}}}}^{\otimes 2}} \\ \mathrm{CUI}(Z^o)_r & \xrightarrow[d_G]{} & (\mathrm{CUI}(Z^o)^{\otimes 2})_r \end{array}$$

Since the corresponding vertical arrows in the left and right squares differ by $\epsilon'$, it follows that the square on the right fails to commute by at most $\epsilon'$. This gives us the inequality of Proposition 3.8.2(g) up to a possible change in $\epsilon'$ stemming from the failure of $\widetilde{{}^{U}I^{F\phi}_{\mathrm{Stdd}}}$ to respect two splittings: the splitting of

$$\tilde{E}_r \subset A(Z^o)_r$$

given by the complementary space $P_r \oplus D_r$ inside the source on the one hand and the splitting of

$$\widetilde{{}^{U}I^{F\phi}_{\mathrm{Stdd}}}(\tilde{E}_r) \subset \prod_{\mathfrak{p}\,|\,p} A(\mathcal{O}_{\mathfrak{p}})$$

induced by the standard basis inside the target on the other hand.[12] This is clearly admissible; we omit the details.

**3.8.13.** Returning to part (e), we must show that our modifications of $\epsilon$ form an algorithmically computable function which goes to zero with $\epsilon$. This is elementary, and fits into the general setting of a valued field $(k, |\cdot|)$ and linear map

$$\phi : k^m \to k^n$$

with kernel $E$. We claim that if

$$|\phi x| < \epsilon$$

then

$$|x - E| < C\epsilon$$

[12] In fact, to decrease the change in $\epsilon$, we could replace the standard basis of $A(\mathcal{O}_{\mathfrak{p}})$ with a basis compatible with the decomposition of the latter into extensions, primitive nonextensions, and decomposables, as we do for $A(Z^o)$. The map $\widetilde{{}^{U}I_{\tilde{\mathcal{A}}}}$ would then be nearly compatible with the splittings, yielding a function which is quadratic in $\epsilon$.

for some algorithmically computable constant $C$. We let $W$ denote the image of $\phi$ and $V$ the coimage, both with induced norms. For $x \in k^m$ we let $\bar{x}$ denote its image in $V$, and we let $\bar{\phi}$ denote the isomorphism

$$V \xrightarrow{\sim} W$$

induced by $\phi$. We fix a metric isomorphism $V = W$ arbitrarily (in practice this would be accomplished by constructing orthonormal bases of both spaces), and we let $C^{-1}$ be the absolute value of the smallest eigenvalue. Then for $x \in k^m$ we have

$$|\phi x| = |\bar{\phi}\bar{x}| \geq C^{-1}|\bar{x}| = C^{-1}|x - E|,$$

independently of the choice of metric isomorphism, which establishes the claim.

**3.8.14.** We turn to part (2) of the proposition: the conditional halting. If the conjectures of Zagier and Goncharov hold for the given input, then our search-space includes an open subscheme $Z^o \subset Z$, a basis $\mathcal{E}_{\leq n}$ of $E_{\leq n}$ consisting of single valued unipotent ($\leq n$)-logarithms which are combinatorially unramified over $Z^o$, and a linearly independent set $\mathcal{P}_{\leq n}$ of unipotent iterated integrals which are combinatorially unramified over $Z^o$ completing $\mathcal{E}_{\leq n} \cup \mathcal{D}_{\leq n}$ to a basis of $A(Z^o)_{\leq n}$. Our claim is that if the Hasse principle holds, then for $\epsilon'$ sufficiently small, the boolean subalgorithm evaluated on the associated datum

$$(Z^o, \mathcal{E}_{\leq n}, \mathcal{P}_{\leq n}, \epsilon')$$

returns *True*. The map

$$\mathfrak{R}_p : \mathbb{Q}_p \otimes \mathrm{Ext}^1_K(\mathbb{Q}(0), \mathbb{Q}(n)) \to \prod_{\mathfrak{p} \mid p} \mathrm{Ext}^1_{\mathcal{O}_\mathfrak{p}}(\mathbb{Q}_p(0), \mathbb{Q}_p(n))$$

from the global motivic Ext group to the product of filtered $\phi$ Ext groups corresponds to the localization map of the Hasse principle through the $p$-adic regulator isomorphism of Soulé [1981] on the source and through the Bloch–Kato exponential map [Bloch and Kato 1990] on the target. So under the Hasse principle, the map

$$\mathrm{Re}_p : A(Z^o)_n \to \prod_{\mathfrak{p} \mid p} A(\mathcal{O}_\mathfrak{p})_n$$

(for $n \geq 2$) is injective near the extension space $E_n$. For any set of linearly independent vectors in a (finite dimensional) normed vector space, there exists an $\epsilon$ such that any set of $\epsilon$-approximations is $\epsilon$-linearly independent. So the claim follows. This concludes the proof Proposition 3.8.2.

**Remark 3.8.15.** Recall that the iterated integrals through which we search to form the set $\mathcal{P}_{\leq n}$ are parametrized by families of sections of $\mathbb{P}^1$ over $Z^o$ (and have no particular relationship to $\mathbb{P}^1 \setminus \{0, 1, \infty\}$). Moreover, the Goncharov exhaustion conjecture places no bound on the height of points needed to form a basis. Thus, the search, as formulated here, is huge and unwieldy. Nevertheless, as a second attempt to convince the reader of the termination, we note that no matter how we order this countably huge set through which we search, if, as predicted by the conjectures, a successful candidate exists, it will, in due course, be met. In practice, we would probably place a height bound $B$ on points and a bound $C$ on the

size of $Z \setminus Z^o$ and then, in turns, increase $B$, decrease $\epsilon'$, increase $C$. Beyond that, as mentioned in the introduction, pairing down the data and ordering it for an efficient search presents an interesting problem in its own right.

### 3.9. *Change of basis algorithm.*

**3.9.1.** We now construct an algorithm which changes the basis constructed by the basis algorithm to one which is compatible with the coproduct up to possible errors of size $\epsilon$. This algorithm takes as input a datum $(Z^o, \tilde{\mathcal{E}}_{\leq n}, \mathcal{P}_{\leq n}, \epsilon, \langle \cdot, \cdot \rangle_\epsilon)$ as in the output of the basis algorithm $A_{\mathrm{Basis}}$, and outputs for each $n' \leq n$, a square matrix of size $a_{n'} \times a_{n'}$ over $\mathbb{Q}$. We denote this algorithm by $A_{\mathrm{Change}}$ and the resulting $n'$-th matrix by

$$A_{\mathrm{Change}}(Z^o, \tilde{\mathcal{E}}_{\leq n}, \mathcal{P}_{\leq n}, \epsilon, \langle \cdot, \cdot \rangle_\epsilon, n').$$

**3.9.2.** For each $n' \leq n$ we fix a set

$$\Sigma^o_{n'} = \{\sigma_{-n',1}, \ldots, \sigma_{-n',e_{n'}}\}$$

of symbols, and define the half-weight of $\sigma_{-n',i}$ to be $-n'$. We may then speak about words in $\Sigma^o_{\geq -n}$ and about the half-weight of a word. Entries in our matrix will be indexed by pairs $(I, w)$, with

$$I \in \tilde{\mathcal{A}}_{n'} = \tilde{\mathcal{E}}_{n'} \cup \mathcal{P}_{n'} \cup \tilde{\mathcal{D}}_{n'}$$

($\tilde{\mathcal{D}}_n$ denoting the set of monomials in $\tilde{\mathcal{B}}_{<n} = \tilde{\mathcal{E}}_{<n} \cup \mathcal{P}_{<n}$ of half-weight $n$ as usual), and $w$ a word in $\Sigma^o$ of half weight $-n'$. We construct the associated matrix entry $a_{I,w}$ by recursion on the length of $w$. We write $\tilde{\mathcal{E}}_{n'}$ as a vector

$$\tilde{\mathcal{E}}_{n'} = (\tilde{\mathcal{E}}_{n',1}, \ldots, \tilde{\mathcal{E}}_{n',e_{n'}})$$

(so $\tilde{\mathcal{E}}_{n',i} = \mathrm{Li}^{U,sv}_{n'}(a_{n',i})$ is a single-valued unipotent $n'$-logarithm). When $w = \sigma_{-n',i}$ is a one-letter word, we set

$$a_{I,w} = \begin{cases} 1 & \text{if } I = \tilde{\mathcal{E}}_{n',i}, \\ 0 & \text{otherwise.} \end{cases}$$

Now suppose $n' = l + m$ with $l, m > 0$, and let $w$ be a word of half-weight $-m$. Using the Goncharov coproduct and the inner product, we expand

$$d_{l,m} I = \sum c_{j,k} \tilde{\mathcal{A}}_{l,j} \otimes \tilde{\mathcal{A}}_{m,k}$$

in the basis

$$\tilde{\mathcal{A}}_l \otimes \tilde{\mathcal{A}}_m = \{\tilde{\mathcal{A}}_{l,j} \otimes \tilde{\mathcal{A}}_{m,k}\}_{j,k}$$

of $A_l \otimes A_m$ to precision $\epsilon$. In terms of the $c_{j,k}$, we define

$$a_{I,\sigma_{-l,i} \cdot w} = \sum_{j,k} c_{j,k} \cdot a_{\tilde{\mathcal{A}}_{l,j}, \sigma_{-l,i}} \cdot a_{\tilde{\mathcal{A}}_{m,k}, w}$$

which equals

$$\sum_k c_{i,k} a_{\tilde{\mathcal{A}}_{m,k},w}$$

if we number the basis $\tilde{\mathcal{A}}_m$ in such a way that

$$\tilde{\mathcal{A}}_{m,j} = \tilde{\mathcal{E}}_{m,j}$$

for $j \in [1, e_m]$.

**3.9.3.** We now make the meaning of the output precise. Let $\mathcal{E}_{n,i}$ denote the projection of $\tilde{\mathcal{E}}_{n,i}$ onto the space $E_n$ of extensions (so in the context of the basis algorithm, we have $\mathcal{E}_{n,i} = \mathrm{Li}_n^{E,sv}(a_{n,i})$). For each $n$, we let $\mathcal{B}_n = \mathcal{E}_n \cup \mathcal{P}_n$, and we let $\mathcal{D}_n$ denote the set of monomials in $\mathcal{B}_{<n}$. According to Proposition 3.2.2,

$$\mathcal{A}_n := \mathcal{B}_n \cup \mathcal{D}_n$$

forms a basis of $A(Z^o)_n$. Let $\sigma_{-n,i} \in \mathcal{U}(Z^o)_{-n}$ denote the element dual to $\mathcal{E}_{n,i}$ relative to this basis. With this interpretation of the set $\Sigma^o$ of symbols $\sigma_{-n,i}$, according to Proposition 3.2.3, $\Sigma^o$ becomes a set of free generators of the free prounipotent group $U(Z^o)$. For $w \in \mathcal{U}(Z^o)_{-n}$ a word in $\Sigma^o$ of half-weight $-n$, let $f_w \in A(Z^o)_n$ denote the corresponding function.

Let $(b_{w,I})_{w,I}$ denote the inverse of the matrix constructed in the algorithm. For $w$ a word of half-weight $-n$, define $\tilde{f}_w \in A(Z^o)_n$ by

$$\tilde{f}_w = \sum_{I \in \tilde{\mathcal{A}}_n} b_{w,I} I.$$

**Proposition 3.9.4.** *In the situation and the notation above, we have*

$$|f_w - \tilde{f}_w| < \epsilon.$$

*Proof.* This is equivalent (up to an admissible change in $\epsilon$) to the estimate

$$|a_{I,w} - \langle w, I \rangle)| < \epsilon.$$

Our algorithm is based on the following two properties of the numbers $\langle w, I \rangle$ for $w$ a word in our set of abstract generators $\Sigma$, and $I$ an element of our concrete basis $\tilde{\mathcal{A}}_n$:

(1) We have

$$\left| \langle \sigma_{n,i}, I \rangle - \begin{cases} 1 & \text{if } I = \tilde{\mathcal{E}}_{n,i}, \\ 0 & \text{otherwise} \end{cases} \right| < \epsilon.$$

(2) We have

$$|\langle \sigma_{n,i} \cdot w, I \rangle - \langle \sigma_{n,i} \otimes w, dI \rangle| < \epsilon.$$

The proposition follows. □

## 4. Construction of geometric algorithms

**4.1.** ***Cocycle-evaluation-image algorithm.*** Fix finite sets $\Sigma_{-1}$, $\Sigma_{-2}$, $\Sigma_{-3}, \ldots, \Sigma_{-n}$ and $\Sigma^\circ_{-1}$, $\Sigma^\circ_{-2}$, $\Sigma^\circ_{-3}, \ldots, \Sigma^\circ_{-n}$ with

$$\Sigma_{-1} \subset \Sigma^\circ_{-1}$$

and $\Sigma^\circ_i = \Sigma_i$ for $i \leq -2$. Set

$$\Sigma = \bigcup \Sigma_i, \quad \Sigma^\circ = \bigcup \Sigma^\circ_i.$$

Let $\mathfrak{n}(\Sigma)$, $\mathfrak{n}(\Sigma^\circ)$ denote the free graded pronilpotent Lie algebras on generators $\Sigma$, $\Sigma^\circ$. As usual, we refer to the grading as the *half-weight*. Let $\mathfrak{n}^{\mathrm{PL}}$ denote the polylogarithmic Lie algebra over $\mathbb{Q}$,

$$\mathfrak{n}^{\mathrm{PL}} = \mathbb{Q}(1) \ltimes \prod_{i=1}^{\infty} \mathbb{Q}(i)$$

with $\mathbb{Q}(i)$ in half-weight $-i$. We write $\mathrm{Hom}^{\mathbb{G}_m}_{\mathrm{Lie}}$ for homogeneous Lie-algebra homomorphisms of graded degree 0. Let $\phi$ denote the natural quotient map

$$\phi : \mathfrak{n}(\Sigma^\circ) \twoheadrightarrow \mathfrak{n}(\Sigma)$$

and let $\mathfrak{ev}$ denote the map

$$\mathrm{Hom}^{\mathbb{G}_m}_{\mathrm{Lie}}(\mathfrak{n}(\Sigma), \mathfrak{n}^{\mathrm{PL}}) \times \mathfrak{n}(\Sigma^\circ)_{\geq -n} \to \mathfrak{n}^{\mathrm{PL}}_{\geq -n} \times \mathfrak{n}(\Sigma^\circ)_{\geq -n}$$

given by

$$\mathfrak{ev}(\mathcal{C}, F) = (\mathcal{C}(\phi(F)), F).$$

Then $\mathfrak{ev}$ is in an obvious sense a map of finite dimensional affine spaces, and it is straightforward to construct an algorithm which computes its scheme-theoretic image. (At the very least, this would be a standard application of elimination theory, but in fact, it should be possible to obtain a closed formula.) We omit the details. We refer to this as the *cocycle-evaluation-image* algorithm. We denote it by $\mathcal{A}_{\mathrm{Eval}}$, and its output, a finite list of elements of

$$S^\bullet(\mathfrak{n}^{\mathrm{PL}}_{\geq -n} \times \mathfrak{n}(\Sigma^\circ)_{\geq -n})^\vee,$$

by $\mathcal{A}_{\mathrm{Eval}}(\Sigma, \Sigma^\circ, n)$.

**4.2.** ***Chabauty–Kim-loci algorithm.***

**4.2.1.** We now construct the Chabauty–Kim-loci algorithm discussed in the introduction. As input it takes an open integer scheme $Z$, a prime $p$ of $\mathbb{Z}$, a natural number $n$, and an $\epsilon \in p^{\mathbb{Z}}$. As output it returns a finite family

$$\tilde{\mathcal{B}} = \tilde{\mathcal{E}} \cup \mathcal{P}$$

of unipotent iterated integrals, and a finite family $\{\tilde{F}_i\}_i$ of elements of the polynomial ring

$$\mathbb{Q}[\tilde{\mathcal{B}}, \log, \mathrm{Li}_1, \mathrm{Li}_2, \ldots, \mathrm{Li}_n],$$

which we denote by $\mathcal{A}_{\mathrm{Loci}}(Z, p, n, \epsilon)$.

**4.2.2.** We run $A_{\mathrm{Basis}}(Z, p, n, \epsilon)$. This gives us our set $\tilde{\mathcal{B}} = \tilde{\mathcal{B}}_{\leq n}$ of unipotent iterated integrals. We run $\mathcal{A}_{\mathrm{Change}}$ on $A_{\mathrm{Basis}}(Z, p, n, \epsilon)$ to obtain a matrix

$$M_{\leq n} = \bigoplus_{i=0}^{n} M_i.$$

**4.2.3.** We let $\Sigma_{-1}$ denote a set of size $e_1 = \dim \mathcal{O}_Z^* \otimes \mathbb{Q}$, $\Sigma_{-1}^o$ a set containing $\Sigma_{-1}$ of size $e_1^o = \dim \mathcal{O}_{Z^o}^* \otimes \mathbb{Q}$, and for each $i \in [2, n]$, $\Sigma_{-i} = \Sigma_{-i}^o$ a set of size

$$e_i = \dim \mathrm{Ext}^1_K(\mathbb{Q}(0), \mathbb{Q}(i)).$$

We run $\mathcal{A}_{\mathrm{Eval}}(\Sigma, \Sigma^o, n)$ to obtain a finite family $\{F_i^{\mathrm{abs}}\}_i$ of elements of $S^\bullet(\mathfrak{n}^{\mathrm{PL}}_{\geq -n} \times \mathfrak{n}(\Sigma^o)_{\geq -n})^\vee$.

**4.2.4.** We pull back along the quotient map

$$\mathfrak{n}(\Sigma^o) \twoheadrightarrow \mathfrak{n}(\Sigma^o)_{\geq -n}.$$

We pull back further along the logarithm

$$U(\Sigma^o) \to \mathfrak{n}(\Sigma^o).$$

Denoting the natural coordinates on $\mathfrak{n}^{\mathrm{PL}}$ by $\log, \mathrm{Li}_1, \mathrm{Li}_2, \mathrm{Li}_3, \ldots$, we obtain a finite family of elements of

$$S^\bullet \mathfrak{n}^{\mathrm{PL}}_{\geq -n} \otimes A(\Sigma^o) = A(\Sigma^o)[\log, \mathrm{Li}_1, \ldots, \mathrm{Li}_n]$$

which are contained in degrees $\leq n$.

**4.2.5.** The matrix $M_{\leq n}$ defines a linear bijection

$$A(\Sigma^o)_{\leq n} \xrightarrow{\sim} \mathbb{Q}[\tilde{\mathcal{B}}]_{\leq n}$$

which we use to obtain the hoped-for family $\{\tilde{F}_i\}_i$. This completes the construction of the algorithm.

**4.2.6.** *Proof of Theorem 2.4.1.* We have a sequence of maps

$$U(X)^{\mathrm{PL}}_{\geq -n, \mathbb{Q}_p} \to \mathfrak{n}(X)^{\mathrm{PL}}_{\geq -n} \times \mathrm{Spec}\, \mathbb{Q}_p \to \mathfrak{n}(X)^{\mathrm{PL}}_{\geq -n} \times \mathfrak{n}(Z^o) \twoheadrightarrow \mathfrak{n}(X)^{\mathrm{PL}}_{\geq -n} \times \mathfrak{n}(Z^o)_{\geq -n}$$

and an associated sequence of Cartesian squares:

$$\begin{array}{ccc}
\mathrm{Hom}^{\mathbb{G}_m}_{\mathrm{Lie}}(\mathfrak{n}(Z), \mathfrak{n}(X)^{\mathrm{PL}}_{\geq -n}) \times \mathfrak{n}(Z^o)_{\geq -n} & \xrightarrow{\overline{\mathfrak{ev}}_n} & \mathfrak{n}(X)^{\mathrm{PL}}_{\geq -n} \times \mathfrak{n}(Z^o)_{\geq -n} \\
\uparrow & & \uparrow \\
\mathrm{Hom}^{\mathbb{G}_m}_{\mathrm{Lie}}(\mathfrak{n}(Z), \mathfrak{n}(X)^{\mathrm{PL}}_{\geq -n}) \times \mathfrak{n}(Z^o) & \xrightarrow{\mathfrak{ev}_n} & \mathfrak{n}(X)^{\mathrm{PL}}_{\geq -n} \times \mathfrak{n}(Z^o) \\
\uparrow & & \uparrow \\
\mathrm{Hom}^{\mathbb{G}_m}_{\mathrm{Lie}}(\mathfrak{n}(Z), \mathfrak{n}(X)^{\mathrm{PL}}_{\geq -n}) \times \mathrm{Spec}\, \mathbb{Q}_p & \xrightarrow{\mathfrak{ev}_n(v_{\mathfrak{p}})} & \mathfrak{n}(X)^{\mathrm{PL}}_{\geq -n} \times \mathrm{Spec}\, \mathbb{Q}_p \\
\uparrow & & \uparrow \\
\mathrm{Hom}^{\mathbb{G}_m}_{\mathrm{Groups}}(U(Z), U(X)^{\mathrm{PL}}_{\geq -n})_{\mathbb{Q}_p} & \xrightarrow{ev_n(u_{\mathfrak{p}})} & U(X)^{\mathrm{PL}}_{\geq -n, \mathbb{Q}_p}
\end{array} \qquad \Gamma$$

We write Im for scheme-theoretic image. By the mere commutativity of the outer square we have a containment of closed subschemes

$$\mathrm{Im}\, ev_n(u_{\mathfrak{p}}) \subset \Gamma^{-1}\, \mathrm{Im}\, \overline{\mathfrak{ev}}_n.$$

In view of Propositions 3.8.2 and 3.9.4, this establishes part 1 of the theorem (correctness of the output upon halting).

For the conditional halting, we contend that formation of scheme-theoretic image along each horizontal map is compatible with pullback along each vertical map. We start at the top. For $v \in \mathfrak{n}(Z^o)$ mapping to $\bar{v} \in \mathfrak{n}(Z)$ and

$$\phi : \mathfrak{n}(Z) \to \mathfrak{n}(X)^{\mathrm{PL}}_{\geq -n}$$

a $\mathbb{G}_m$-equivariant homomorphism, $\phi(\bar{v})$ depends only on the image of $v$ in $\mathfrak{n}(Z^o)_{\geq -n}$. So

$$\mathrm{Im}\, \mathfrak{ev}_n = (\mathrm{Im}\, \overline{\mathfrak{ev}}_n) \times \mathfrak{n}(Z^o)_{<-n}.$$

We go on to the middle square. By the period conjecture, $A(Z^o) \to \mathbb{Q}_p$ is flat. (In general, if $A$ is integral, $K$ a field, and $\psi : A \to K$ is injective, then $\psi$ is flat, since it factors as a localization map followed by a field extension.) Hence formation of the scheme-image is compatible with pullback. Turning to the bottom square, the vertical maps are iso, so this is clear. This completes the proof of Theorem 2.4.1.

## 5. Construction of analytic algorithm

**5.1.** We now construct an algorithm for deciding whether the number of zeroes of a $p$-adic power series in a given ball is zero or one, given a sufficiently close approximation. This is quite standard, but we are unaware of a reference for this exact problem.

More precisely, we construct a boolean-valued algorithm which takes as input a boolean $b \in \{0, 1\}$, natural numbers $N$, $r$ and $h$, and a polynomial with rational coefficients

$$\tilde{F} = \sum_{i=0}^{N} \tilde{a}_i T^i$$

which has at most $b$ ($\mathbb{Q}_p$-rational) roots inside the disk of radius $p^{-r}$. We call this algorithm the *root criterion algorithm* and denote the output by $\mathcal{A}_{\mathrm{RC}}(b, N, r, h, \tilde{F})$. We first announce the meaning of the output as a remark.

**Remark 5.2.** In our application below,

$$F = \sum_{i=1}^{\infty} a_i T^i$$

will be a power-series expansion of a polynomial in logarithms and polylogarithms of half-weight $h$ over $\mathbb{Q}_p$, and $\tilde{F}$ will be an approximation of $F$ with arithmetic precision $p^{-r}$ and geometric precision $e^{-N}$. Suppose $\mathcal{A}_{\mathrm{RC}}(b, N, r, h, \tilde{F}) = \mathit{True}$ . Then $F$ has at most $b$ roots within the disk of radius $p^{-r}$. This amounts to an elementary use of Newton polygons, together with the growth estimate

$$v(a_k) \geq -h \log_p(k)$$

which follows from Proposition 6.7 of Besser and de Jeu [2008].

Case 1: $b = 0$.

(1) If $\tilde{a}_0 = 0$, return *False*.

(2) Compute real solutions to

$$v(\tilde{a}_0) - rt = -h \log_p t$$

to within 0.5. If there are none, return *True*.

(3) Otherwise, there are two solutions $t_L < t_R$. If $t_R > N$, return *False*.

(4) Check the condition

$$v(\tilde{a}_k) > v(\tilde{a}_0) - rk$$

for $1 \leq k \leq t_R$. If the condition holds, return *True*. Otherwise return *False*.

Case 2.1: $b = 1$, $\tilde{a}_0 = 0$.

(1) If $\tilde{a}_1 = 0$, return *False*.

(2) Compute solutions to

$$-r(t-1) + v(\tilde{a}_1) = -h \log_p t$$

to within 0.5. If there are none, return *True*.

(3) Otherwise, there are two solutions $t_L < t_R$. If $t_R > N$, return *False*.

(4) Check the condition

$$v(\tilde{a}_k) > -r(k-1) + v(\tilde{a}_1)$$

for $2 \leq k \leq t_R$. If this condition holds, return *True*. Otherwise, return *False*.

Case 2.2: $b = 1$, $\tilde{a}_0 \neq 0$. In this case, we have

$$v(\tilde{a}_1) = v(\tilde{a}_0) - r \tag{$*$}$$

and

$$v(\tilde{a}_0) - rt = -h \log_p t \tag{$**$}$$

has two solutions $t_L < t_R$:

(1) If $t_R > N$, return *False*.

(2) Check the condition

$$v(\tilde{a}_i) > v(\tilde{a}_0) - ri$$

for $2 \leq i \leq t_R$. If this condition holds, return *True*, otherwise return *False*.

## 6. Numerical approximation

We present the results of the unpublished work [Dan-Cohen and Chatzistamatiou 2014] concerning computation of $p$-adic iterated integrals on the projective line. For background we refer the reader to [Furusho 2004; 2007] and to [Chatzistamatiou 2017]. The results of this section should be compared with prior and with concurrent works by Besser and de Jeu [2008], by Jarossay [2016], and by Ünver [2019].

**6.1.** Let $K$ denote a finite extension of $\mathbb{Q}_p$. A general p-adic iterated integral on the line with arbitrary punctures with good reduction is a multiple polylogarithm up to sign: if the points $a_1, \ldots, a_m$ of $\mathcal{O}_K^*$ lie in distinct residue disks, then we have

$$\int_{1_0}^{a_{m+1}} \left(\frac{dt}{t}\right)^{n_m-1} \frac{dt}{t-a_m} \cdots \left(\frac{dt}{t}\right)^{n_1-1} \frac{dt}{t-a_1} = (-1)^m \operatorname{Li}_{n_1,\ldots,n_m}\left(\frac{a_2}{a_1}, \frac{a_3}{a_2}, \ldots, \frac{a_{m+1}}{a_m}\right)$$

(see Theorem 2.2 of [Goncharov 2001]). Here $1_0$ denotes the tangent vector 1 at 0. We will restrict attention to the case $m = 1$, $a_1 = 1$, i.e., to iterated integrals on $\mathbb{P}^1 \setminus \{0, 1, \infty\}$; in terms of multiple polylogarithms, this means restricting attention to multiple polylogarithms of one variable. We do this merely for concreteness: the difficulty in going from this case to the general case is largely a notational one.

**6.2.** We work with the rigid analytic space equal to the closed unit disk minus the residue disk about 1:

$$U = \operatorname{Spm} K\{t, u\}/(u(t-1) - 1),$$

with log structure induced by the divisor $\{0\}$.[13] The coordinate ring $\mathcal{O}(U)$ has a map

$$\rho_0 : \mathcal{O}(U) \to B := K\{t\}$$

as well as a map

$$\rho_1 : \mathcal{O}(U) \to A := K\{w, u\}/(wu - 1)$$

sending

$$t \mapsto w + 1$$

(recentering the unit circle). Both $\rho_0$ and $\rho_1$ are injective, and on a practical level, all computations that occur within the algorithm may in fact be carried out inside the rings $A$ and $B$ up to given arithmetic and geometric precision. There will be no need to verify convergence (or overconvergence) algorithmically. As usual, we do not keep track of the error incurred.

**6.3.** ***Remark.*** There are several possible treatments of the singular points $0, 1, \infty$. In one approach, we puncture over the residue field $k$ (which corresponds, via Berthelot's theory of rigid analytic tubes, to removing residue disks over $\operatorname{Spm} K$). In another approach, we include the points $0, 1, \infty$, allowing log poles instead of puncturing; in Berthelot's construction of isocrystals, this does not require the use of log geometry. This approach is convenient for dealing with tangential base-points. (There is also the possibility of working with the rigid analytic thrice punctured line $\mathbb{P}^1_{\operatorname{Spm} K} \setminus \{0, 1, \infty\}$, needed when considering points of bad reduction.) The resulting categories of unipotent isocrystals on the special fiber or of unipotent connections on any of the above rigid analytic spaces are all canonically equivalent. Our use of the rigid analytic space $U$ means that we choose to mix the first two approaches.

**6.4.** We give a brief sketch of the definition of the p-adic multiple polylogarithms we wish to evaluate. We refer to Besser [2012; 2002] and to Chatzistamatiou [2017] for more thorough accounts.

We let $\mathrm{unVIC}(U)$ denote the category of unipotent vector bundles with integrable connection with log poles at 0. Its equivalence with, say, the category of unipotent isocrystals on $\mathbb{P}^1_k \log 0 \setminus \{1, \infty\}$ endows it with a Frobenius pullback functor

$$F^* : \mathrm{unVIC}(U) \to \mathrm{unVIC}(U).$$

If $\omega$, $\omega'$ are fiber functors *compatible with Frobenius*, then $F^*$ acts on the Tannakian path torsor

$${}_{\omega'}P_{\omega} = \mathrm{Isom}^{\otimes}(\omega, \omega').$$

The exact notion of compatibility is a bit subtle, and we refer the reader to [Chatzistamatiou 2017] for a detailed discussion. The result of this detailed discussion however, is that in our setting, a naive approach will be sufficient.

[13] For us, this merely means that we will be working with differential forms with log poles at the origin; we will not make use of log geometry.

We consider the $K$-rational fiber functors

$$\mathrm{unVIC}(U) \overset{\omega_{1_0}}{\underset{\omega_x}{\rightrightarrows}} \mathrm{Vect}\, K$$

associated to the tangent vector $1_0$ and to $x$ respectively. According to Besser [2012], there's a unique Frobenius fixed point $p^{\mathrm{cris}} \in ({}_x P_{1_0})(K)$.

We let $\tilde{E}$ denote the KZ-connection over $U$: $\tilde{E} = \mathcal{O}_U \langle\langle e_0, e_1 \rangle\rangle$, with connection given by

$$\nabla(W) = e_0 W \frac{dt}{t} + e_1 W \frac{dt}{1-t}$$

for any word $W$ in $e_0, e_1$. Thus, if

$$\mathcal{L} = \sum_W \mathcal{L}_W W$$

is an arbitrary section, we have

$$\nabla(W) = d\mathcal{L} + e_0 \mathcal{L}\omega_0 + e_1 \mathcal{L}\omega_1$$

where

$$\omega_0 := \frac{dt}{t}, \quad \omega_1 := \frac{dt}{t-1}, \quad \text{and} \quad d\mathcal{L} = \sum_W \mathcal{L}'_W W dt$$

is given by differentiating the coefficient functions.

Each $K$-rational fiber of $\tilde{E}$ is canonically equal to

$$\mathcal{U} := K \langle\langle e_0, e_1 \rangle\rangle;$$

let $\lambda_W$ denote the linear functional

$$\mathcal{U} \to K$$

associated to the word $W$. The Tannakian path $p^{\mathrm{cris}}$ gives rise to a linear map

$$p^{\mathrm{cris}}_{\tilde{E}} : \tilde{E}(1_0) \to \tilde{E}(x).$$

In terms of the path $p^{\mathrm{cris}}$ and the linear functional $\lambda_W$, we define the p-adic multiple polylogarithms we wish to evaluate by

$$\mathrm{Li}_W(x) = \lambda_W(p^{\mathrm{cris}}_{\tilde{E}}(1)).$$

**6.5.** Our algorithm depends on a notion of *residue over residue disk*. In turn, this depends on an elementary lemma:

**Lemma.** *Suppose we have a congruence relation*

$$\sum_{i,j \geq 0} a_{i,j} w^i u^j \equiv \sum_{i,j \geq 0} b_{i,j} w^i u^j \mod (wu - 1)$$

*in the restricted formal power series ring $K\{w, u\}$. Then we have an equality of convergent sums*

$$\sum_{j=0}^{\infty} a_{j+1,j} = \sum_{j=0}^{\infty} b_{j+1,j}$$

*in the nonarchimedean field $K$.*

*Proof.* We provide the details of this elementary verification. In a nonarchimedean field, a sum whose terms tend to zero converges, so the convergence is immediate. For the equality, suppose the congruence relation is witnessed by the equation

$$\sum_{i,j\geq 0} (a_{i,j} - b_{i,j}) = (wu - 1) \sum_{i,j\geq 0} c_{i,j} w^i u^j.$$

We then have for each $n, k \in \mathbb{N}$,

$$c_{n+k+1,k+1} = \sum_{r=1}^{k} a_{n+r,r} - \sum_{r=1}^{k} b_{n+r,r}.$$

Restricting attention to $n = 1$ and taking the limit as $k \to \infty$, we obtain the result. □

Returning to our definition of the residue, we consider the space

$$\Omega^1(U) = \mathcal{O}(U)\frac{dt}{t}$$

of rigid analytic 1-forms with log poles at the origin. The pullback of an arbitrary element

$$\omega = f(t, u)\frac{dt}{t}$$

along the map $\rho_1$ defined in segment 6.2 is given by

$$\rho_1^* \omega = (w + 1)^{-1} f(w + 1, w^{-1}) dw.$$

We define

$$\mathrm{Res}_{D(1)}\, \omega := \mathrm{Res}_{w<1} (w + 1)^{-1} f(w + 1, w^{-1}),$$

where

$$\mathrm{Res}_{w<1} : K\{w, u\}/(wu - 1) \to K$$

is defined by

$$\mathrm{Res}_{w<1}\left( \sum_{i,j} a_{i,j} w^i u^j \right) := \sum_{j=0}^{\infty} a_{j+1,j}.$$

**6.6.** We will construct elements $\tau_1 \in \mathcal{U}$ and $L \in \tilde{E}(U)$ recursively. To do so, we set

$$\omega_1' := \frac{pt^{p-1}dt}{t^p - 1}.$$

We denote the length of a word $W$ by $|W|$. We also let $|W|_i$ denote the number of occurrences of the letter $e_i$. We set

$$\tau_1 := pe_1 + \sum_{\substack{|W|\geq 2 \\ |W|_1 \geq 1}} \tau_W W \tag{♯}$$

and

$$L = 1 + \sum_{|W|\geq 1} L_W W. \tag{♮}$$

We require that $L$, which will be determined by a system of differential equations, satisfy the initial condition

$$L_W(0) = 0 \tag{♭}$$

for all nonempty words $W$. We construct the functions $L_W$ as elements of the ring $A$. There, we set $t := w + 1$. We may transport the 1-forms $\omega_0$, $\omega_1$, $\omega_1'$ to $A$ in the obvious way. We may also formally differentiate as well as take residues about the open disk $|w| < 1$, which we continue to denote by $\mathrm{Res}_{D(1)}$. A 1-form $\omega$ in the free rank-one module $Adw$ has a primitive in $A$ if and only if $\mathrm{Res}_{D(1)}\,\omega = 0$. With these notations, and this last fact in mind, we define $\tau_W$ in terms of lower terms by

$$\tau_W = -\mathrm{Res}_{D(1)}\left(p(L_{W/e_0} - L_{e_0\backslash W})\omega_0 + \sum_{\substack{W=W'W'' \\ W',W''\neq\varnothing}} L_{W'}\tau_{W''}\omega_1 - L_{e_1\backslash W}\omega_1'\right). \tag{∗}$$

Notationally, the term with the left-division $e_1\backslash W$ is equal to 0 if $W$ is not left-divisible by $e_1$. We define $L_W$ in terms of lower terms by the differential equation

$$dL_W = p(L_{W/e_0} - L_{e_0\backslash W})\omega_0 + \sum_{\substack{W=W'W'' \\ W''\neq\varnothing}} L_{W'}\tau_{W''}\omega_1 - L_{e_1\backslash W}\omega_1', \tag{⋆}$$

the equation being solvable over $A$ since equation (∗) above guarantees that the right-hand side has no residue. Again, terms with an impossible left or right division are to be interpreted as zero.

**6.7.** We now use the elements $\tau_W \in K$ to construct more elements $\tau_W^V \in K$ indexed by pairs of words $V$, $W$ with

$$V \subset W$$

by which we mean that $V$ occurs as an ordered subsequence of (not necessarily contiguous) letters. To do so, we let $\tau$ be the endomorphism of $\mathcal{U}$ determined by

$$\tau(e_0) = pe_0 \tag{♯}$$

and

$$\tau(e_1) = \tau_1. \tag{♮}$$

We then define $\tau_W^V$ by

$$\tau(V) = \sum_{W \supset V} \tau_W^V W. \tag{♭}$$

**6.8.** We construct certain rational numbers $c_{i,j,W}$ ($i, j \in \mathbb{N}$, $W$ a word in $e_0, e_1$) which arise from the KZ-connection. The $k$-th power of the covariant derivative applied to a word $W'$,

$$\left(\frac{\nabla_{\partial/\partial t}^k W'}{k!}\right)$$

is of the form

$$\sum_{\substack{i+j=k \\ |W|_1 \le j \\ |W|_0 \le i}} \frac{c_{i,j,W}}{t^i (t-1)^j} W W'.$$

To compute the coefficients $c_{i,j,W}$ algorithmically, we simply apply

$$\frac{\nabla_{\partial/\partial t}}{i+j+1}\left(\frac{1}{t^i(t-1)^j} W'\right) = \frac{1}{i+j+1}\left(\frac{-i}{t^{i+1}(t-1)^j} I + \frac{-j}{t^i(t-1)^{j+1}} I + \frac{1}{t^{i+1}(t-1)^j} e_0 + \frac{0}{t^i(t-1)^{j+1}} e_1\right) W'$$

iteratively and collect terms.

**6.9.** All ingredients above are independent of the endpoint. We now fix the point $x \in \mathbb{P}^1 \setminus \{0, 1, \infty\}(\mathcal{O}_K)$ at which we wish to evaluate. In terms of the rational coefficients constructed in segment 6.8, and in terms of a lift $\sigma$ of Frobenius to $K$, we define for each word $W$, an element $\epsilon_W(x) \in K$ by the infinite sum

$$\epsilon_W(x) = \sum_{\{i,j \in \mathbb{N} | i \ge |W|_0, j \ge |W|_1\}} \frac{c_{i,j,W} (x^p - x)^{i+j}}{(x^\sigma)^i (x^\sigma - 1)^j}.$$

We will discuss its convergence below.

**Theorem 6.9.1.** (*Joint with Andre Chatzistamatiou.*) *Let $T$ be a word in $\{e_0, e_1\}$, let $p$ be a prime, let $K$ be a finite extension of $\mathbb{Q}_p$ and let $x \in \mathbb{P}^1 \setminus \{0, 1, \infty\}(\mathcal{O}_K)$. Then the $p$-adic multiple polylogarithm $\mathrm{Li}_T^p(x)$ is given in terms of the values $L_W(x)$, in terms of the constants $\tau_W^V$, in terms of the values $\epsilon_W(x)$, and in terms of multiple polylogarithms of lower weight, by*

$$\mathrm{Li}_T^p(x) = (-1)^{|T|_0}(1 - p^{|T|})^{-1} \sum_{\substack{T = U W' W \\ W \supset V \\ V \ne T}} (-1)^{|V|_0} \epsilon_U(x) L_{W'}(x) \tau_W^V \mathrm{Li}_V^p(x).$$

(*In particular, the terms $\epsilon_U(x)$ and $L_{W'}(x)$ appearing on the right converge.*)

**6.10.** ***Proof of Theorem 6.9.1.***

**6.10.1.** Let $\phi$ denote the Frobenius lift

$$U \to U$$

over $\sigma$ given by $\phi(t) = t^p$ and let $\phi'$ be a lift of Frobenius which fixes the endpoint $x$. For instance, when $K = \mathbb{Q}_p$ we may set $\phi'(t) = (t-x)^p + x$. (Over)convergence gives rise to a canonical isomorphism

$$\epsilon : \phi^* \tilde{E} \xrightarrow{\sim} \phi'^* \tilde{E}.$$

We recall that the pullback along the semilinear Frobenius lift $\phi$ is accomplished in two steps as a base-change along $\sigma$ followed by a $K$-linear pullback, as in the familiar diagram

$$\begin{array}{ccccc} U & \xrightarrow{\phi/K} & U^\sigma & \xrightarrow{\sigma} & U \\ & \searrow & \downarrow & & \downarrow \\ & & \operatorname{Spm} K & \xrightarrow[\sigma]{} & \operatorname{Spm} K \end{array}$$

and similarly for $\phi'$. In concrete terms, Frobenius-invariance of $p^{\mathrm{cris}}$ means that the square

$$\begin{array}{ccc} (\phi^*\tilde{E})(1_0) & \xrightarrow{p^{\mathrm{cris}}_{\phi^*\tilde{E}}} & (\phi^*\tilde{E})(x) \\ \| & & \sim \downarrow \epsilon(x) \\ \sigma^*\tilde{E}(1_0) & \xrightarrow[\sigma^*(p^{\mathrm{cris}}_{\tilde{E}})]{} & \sigma^*\tilde{E}(x) \end{array}$$

commutes.

**6.10.2.** Together with the choice of unit vector $1 \in \tilde{E}(1_0)$, the KZ-connection corepresents the fiber functor: for any unipotent connection $E$, we have

$$\mathrm{Hom}(\tilde{E}, E) = E(1_0).$$

This is proved in a surprising way (using complex iterated integrals) by Kim [2009], and in a more straightforward way (via a certain iterative construction of universal extensions which has appeared in various places in the literature) for instance by Chatzistamatiou [2017]. This gives us a canonical identification between the fiber $\tilde{E}(1_0)$ and the completed universal enveloping algebra of the unipotent fundamental group. It also means that there's a unique overconvergent horizontal morphism

$$\theta : \tilde{E} \to \phi^* \tilde{E}$$

such that

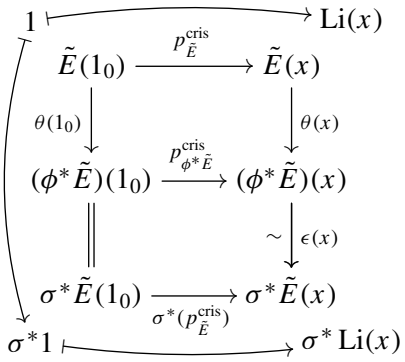
$$\theta(1_0) : \tilde{E}(1_0) \to (\phi^*\tilde{E})(1_0) = \sigma^*(\tilde{E}(1_0))$$

sends $1 \mapsto \sigma^* 1$.

**6.10.3.** Let $\mathrm{Li}(x)$ denote the noncommutative formal power series

$$\mathrm{Li}(x) = \sum_W \mathrm{Li}_W(x) W.$$

Placed side by side, $\theta$, $\epsilon$, and $p^{\mathrm{cris}}$ form a commutative diagram like so:




We will see that $\mathrm{Li}(x)$ is uniquely determined by the equation

$$\sigma^*\,\mathrm{Li}(x) = \epsilon(x)\theta(x)(\mathrm{Li}(x)).$$

**6.10.4.** It follows from the analysis of tangential fiber functors in [Chatzistamatiou 2017, Section 3.5] that there's a canonical isomorphism from the fiber functor $\omega_{1_0}$ to the functor

$$\omega_0 : E \mapsto E(0)$$

which is compatible with the action of our chosen Frobenius lift $\phi$. We claim that the fiber $\theta(0)$ of $\theta$ at 0, regarded as a map

$$\mathcal{U} \to \sigma^*\mathcal{U}$$

is equal to $\tau$, and that the value $\theta(U)(1)$ of $\theta$ at the identity element 1 of $\tilde{E}(U)$ is equal to $L$. These facts follow from two key properties of $\theta$, which in turn follow from a certain functorial characterization of $\theta$. If

$$f : T \to U$$

is a rigid analytic space over $U$, and $E$ is a quasicoherent sheaf over $U$, we set

$$E(T) := \Gamma(T,\, f^*E).$$

We let $\mathrm{Vect}\, T$ denote the category of vector sheaves. We let $\omega_T$ denote the functor

$$\mathrm{Vect}\, T \leftarrow \mathrm{unVIC}(U \log 0)$$

induced by $f$ (forget the connection and pull back). We let $\omega_{0_T}$ denote the composite

$$\mathrm{Vect}\, T \leftarrow \mathrm{Vect}\, K \xleftarrow{\omega_0} \mathrm{unVIC}(U).$$

We let $\phi_* \omega_T$ denote the composite

$$\mathrm{Vect}\, T \xleftarrow{\omega_T} \mathrm{unVIC}(U) \xleftarrow{\phi^*} \mathrm{unVIC}(U),$$

and similarly for $\omega_{0_T}$. Then we have canonical isomorphisms

$$\tilde{E}(T) = \mathrm{Hom}(\omega_{0_T}, \omega_T), \quad \text{and} \quad (\phi^* \tilde{E})(T) = \mathrm{Hom}(\phi_* \omega_{0_T}, \phi_* \omega_T),$$

which are natural in $T$. Moreover, translated through these isomorphisms, $\theta$ sends a 2-morphism

$$\omega_{0_T} \to \omega_T$$

to its composite with the 1-morphism $\phi^*$. It follows, on the one hand, that the fiber $\theta(0)$ is the $\sigma$-linear ring homomorphism induced by the action of Frobenius on the unipotent fundamental group, and on the other hand, that the map of global sections $\theta(U)$ is equivariant for the right-action of $\tilde{E}(0)$ on $\tilde{E}(U)$ and for the right action of $\phi^* \tilde{E}(0)$ on $\phi^* \tilde{E}(U)$ through $\theta(0)$. This last property means that for any $g \in \mathcal{U}$, we have

$$\theta(U)(g) = \theta(U)(I) \cdot \theta(0)(g). \qquad \text{(equivariance)}$$

Thus, $\theta$ is completely determined by two small pieces: a *horizontal* piece $\mathcal{L} := \theta(U)(1)$, and a *vertical* piece $\mathcal{T} := \theta(0)$. Moreover, $\mathcal{T}$ is determined by its action on the generators $e_0$ and $e_1$. Finally,

$$\mathcal{T}(e_0) = p e_0$$

and $\mathcal{T}_1 := \mathcal{T}(e_1)$ has constant term 0 and linear term $p e_1$.

**6.10.5.** We will now show that $L = \mathcal{L}$ and $\tau = \mathcal{T}$. We first obtain the initial condition for $\mathcal{L}$:

$$\mathcal{L}(0) = \theta(U)(1)(0) = \theta(0)(1(0)) = 1.$$

Here "1" denotes the empty word, regarded first as a section of $\tilde{E}$ over $U$, and then as a section of $\phi^* \tilde{E}$ over $U$. Its *value* $1(0)$ is the unit element of the fiber $\tilde{E}(0)$, which gets sent to the unit element of $(\phi^* \tilde{E})(0)$ since $\phi(0)$ is a homomorphism.

By the horizontality of $\theta$, we have the equation

$$\nabla'(\theta(U)(1)) = \theta(U)(\nabla 1) \qquad (\sharp)$$

in the $\mathcal{O}(U)$-module

$$\Gamma(U, (\phi^* \tilde{E}) \otimes \Omega^1_U) = (\phi^* \tilde{E})(U) \otimes \Omega^1(U) = (\phi^* \tilde{E})(U) \frac{dt}{t}.$$

Here, $\nabla'$ denotes the pullback of $\nabla$ along the Frobenius lift $\phi$. By the equivariance property, this equation may be rewritten in terms of $\mathcal{L}$ and $\mathcal{T}$ as follows. We denote the free generators $\phi^* e_i$ of $\phi^* \tilde{E}$ simply by $e_i$. We let $\omega_i'$ denote the pullback of $\omega_i$ by the Frobenius lift $\phi$. With this notation, we have

$$\begin{aligned} d\mathcal{L} + \sum_{i=0,1} e_i \mathcal{L} \omega_i' &= \nabla \mathcal{L} \\ &= \theta(U)(\nabla(1)) \\ &= \theta(U)\left( \sum_{i=0,1} e_i \omega_i \right) \\ &= \sum_{i=0,1} \theta(U)(e_i)\omega_i \\ &= \sum_{i=0,1} \mathcal{L} \cdot \mathcal{T}(e_i(0))\omega_i. \quad \text{(by equivariance)} \end{aligned} \tag{♮}$$

Plugging in $pe_0$ for $\mathcal{T}(e_0)$ and

$$\mathcal{T}_1 = pe_1 + \sum_{|W| \geq 2} \mathcal{T}_W W$$

for $\mathcal{T}(e_1)$, we obtain

$$d\mathcal{L} + pe_0 \mathcal{L}\omega_0 + e_1 \mathcal{L}\omega_1' = p\mathcal{L}e_0\omega_0 + \mathcal{L}\mathcal{T}_1\omega_1. \tag{♭}$$

Modulo the augmentation ideal $I$, (♭) becomes

$$d\mathcal{L}_\varnothing = 0.$$

Hence, from the initial condition, we find that

$$\mathcal{L}_\varnothing = 1.$$

Thus $\mathcal{L}$ has the form

$$\mathcal{L} = 1 + \sum_{|W| \geq 1} \mathcal{L}_W W.$$

Projecting (♭) onto the $W$-coordinate for an arbitrary word $W$, we find that the functions $L_W := \mathcal{L}_W$ satisfy segment 6.6(⋆). Applying $\mathrm{Res}_{D(1)}$ to both sides of segment 6.6(⋆), we find that the constants $\tau_W := \mathcal{T}_W$ satisfy segment 6.6(∗). This completes the verification that $\mathcal{L} = L$ and $\mathcal{T} = \tau$.

An overconvergent 1-form $\omega$ on $U$ with log poles at the origin has an overconvergent primitive if and only if $\mathrm{Res}_0\, \omega = 0$ and $\mathrm{Res}_{D(1)}\, \omega = 0$. It follows that $L$ is overconvergent. In particular, the values $L_W(x)$ converge.

**6.10.6.** To compute $\epsilon$, we write

$$\begin{aligned}\epsilon(\phi^* W') &= \sum_{k=0}^{\infty} (\phi^\sharp t - \phi'^\sharp t)^k \phi'^* \left( \frac{\nabla^k_{\partial/\partial t} W'}{k!} \right) \\ &= \sum_k (\phi^\sharp t - \phi'^\sharp t)^k \cdot \phi'^* \left( \sum_{\substack{i+j=k \\ |W|_0 \le i \\ |W|_1 \le j}} \frac{c_{i,j,W}}{t^i (t-1)^j} W W' \right) \\ &= \sum_{\substack{i,j \in \mathbb{N} \\ |W|_0 \le i \\ |W|_1 \le j}} \frac{c_{i,j,W} (\phi^\sharp t - \phi'^\sharp t)^k}{\phi'^\sharp (t^i (t-1)^j)} \phi'^* (W W').\end{aligned}$$

The coefficient of $\phi'^*(WW')$ is independent of $W'$; setting $\epsilon_W$ equal to this coefficient and valuating at $t = x$ we obtain the formula given in segment 6.9. The prounipotence of $\tilde{E}$ implies convergence of $\epsilon$. In particular, the values $\epsilon_W(x)$ converge.

**6.10.7.** Setting

$$v = \sum_T v_T T,$$

the equation

$$\epsilon(x)\theta(x)(v) = v$$

becomes

$$\begin{aligned}\sum_T v_T T &= \epsilon(x)\theta(x) \left( \sum_V v_V V \right) \\ &= \sum_V v_V \epsilon(x) \left( \sum_{\substack{W' \\ W \supset V}} L_{W'}(x) \tau^V_W W' W \right) \\ &= \sum_V v_V \sum_{\substack{W' \\ W \supset V}} L_{W'}(x) \tau^V_W \sum_U \epsilon_U(x) U W' W \\ &= \sum_{\substack{V,U,W' \\ W \supset V}} \epsilon_U(x) L_{W'}(x) \tau^V_W v_V U W' W \\ &= \sum_T \left( \sum_{\substack{T = UW'W \\ W \supset V}} \epsilon_U(x) L_{W'}(x) \tau^V_W v_V \right) T.\end{aligned}$$

Hence

$$v_T = \sum_{\substack{T = UW'W \\ W \supset V}} \epsilon_U(x) L_{W'}(x) \tau^V_W v_V = p^{|T|} v_T + \sum_{\substack{T = UW'W \\ W \supset V \\ V \neq X}} \epsilon_U(x) L_{W'}(x) \tau^V_W v_V.$$

So the equations

$$v_I = 1 \quad \text{and} \quad \epsilon(x)\theta(x)(v) = v$$

are equivalent to

$$v_I = 1$$

and

$$(1 - p^{|T|})v_T = \sum_{\substack{T=UW'W \\ V \neq X \\ W \supset V}} \epsilon_U(x) L_{W'}(x) \tau_W^V v_V$$

from which we obtain the formula for

$$\mathrm{Li}_T(x) = v_T$$

given in the theorem.[14] This completes the proof of Theorem 6.9.1.

**6.11.** We apply this to the computation of $p$-adic multiple zeta values. We have the path composition formula

$$\int_x^z \omega_n \cdots \omega_1 = \left( \sum_{i=0}^{n} \int_y^z \omega_n \cdots \omega_{i+1} \cdot \int_x^y \omega_i \cdots \omega_1 \right)$$

and the path reversal formula

$$\int_y^x \omega_n \cdots \omega_1 = (-1)^n \int_x^y \omega_n \cdots \omega_1.$$

If $f : Y \to X$ is an isomorphism of rigid curves, then

$$\int_{f(x)}^{f(y)} \omega_n \cdots \omega_1 = \int_x^y (f^*\omega_n) \cdots (f^*\omega_1).$$

Applying this to $\omega_i \in \left\{ \frac{dx}{x}, \frac{dx}{1-x} \right\}$, to

$$x \mapsto 1 - x$$

on $\mathbb{P}^1$, and to an auxiliary point $y$, we obtain

$$\int_y^{\vec{-1}_1} \omega_n \cdots \omega_1 = (-1)^n \int_{\vec{-1}_1}^y \omega_n \cdots \omega_1 = \int_{\vec{1}_0}^{1-y} \omega_n^\circ \cdots \omega_1^\circ$$

where $\left(\frac{dx}{x}\right)^\circ = \left(\frac{dx}{1-x}\right)$ and $\left(\frac{dx}{1-x}\right)^\circ = \left(\frac{dx}{x}\right)$. So

$$\zeta(W) = \int_{\vec{1}_0}^{\vec{-1}_1} \omega_W = \sum_{W=W''W'} \int_y^{\vec{-1}_1} \omega_{W''} \int_{\vec{1}_0}^y \omega_{W'} = \sum_{W=W''W'} \mathrm{Li}_{W''^\circ}(1-y)(\mathrm{Li}_{W'}\, y).$$

[14] Actually, differing sign conventions necessitate the addition of the signs seen in the final formula to accord with the traditional definition.

## 7. The equation-solving algorithm

### 7.1. *Construction of the algorithm.*

**7.1.1.** We now construct the promised algorithm for totally real fields. Our algorithm takes as input an open integer scheme $Z$ and outputs a finite set of elements of $X(Z)$. We denote the output by $\mathcal{A}_{\mathrm{ES}}(Z)$. As explained in the introduction, the success of the algorithm depends on first finding $X(Z)$ by a naive search, and then proving that there are no other points by verifying that $X(Z_{\mathfrak{p}})_n = X(Z)$. Recall also (from our formulation of the convergence conjecture (Conjecture 2.1.4)) that success (i.e., halting) depends also on the absence of repeated roots for $n$ sufficiently large.

**7.1.2.** We find a prime $p$ of $\mathbb{Z}$ in the image of $Z$, for which $Z$ is totally split. We fix arbitrarily a prime $\mathfrak{p}$ of $Z$ lying above $p$.

**7.1.3.** Our algorithm searches through the set of triples $(n, N, \epsilon)$, $n, N \in \mathbb{N}$, $\epsilon$ in a countable subset of $\mathbb{R}_{>0}$ with accumulation point 0. After each attempt, we increase $n$ and $N$ and decrease $\epsilon$. To each such triple, our algorithm assigns a set $X(Z)_n$ of points of $X(Z)$ and a boolean. If the boolean output is *True*, then we output $X(Z)_n$. If the boolean output is *False*, then we continue the search.

To produce the set $X(Z)_n$, we spend $n$ seconds searching for points.[15] To produce the boolean output, we follow the steps described in segements 7.1.4–7.1.9 below.

**7.1.4.** Partition $X(\mathcal{O}_{\mathfrak{p}})$ into $\epsilon$-balls, decreasing $\epsilon$ as needed to ensure that each ball contains at most one element of the set $X(Z)_n$ (our, potentially incomplete, list of integral points).

**7.1.5.** Run $\mathcal{A}_{\mathrm{Loci}}(Z, p, n, \epsilon)$ to obtain a family $\{\tilde{F}_i\}_i$ of polylogarithmic functions. Symmetrize the family with respect to the $S_3$-action using the formulas 2.1.3(∗). Set $h_i$ equal to the half-weight of $\tilde{F}_i$.

**7.1.6.** We focus our attention on an $\epsilon$-ball $B$ containing a rational representative $y \in B$. Expand each polylogarithmic function $\tilde{F}_i$ to arithmetic precision $\epsilon$ and geometric precision $e^{-N}$ about $y$; denote the result by $\widetilde{\tilde{F}}_i^{\mathfrak{p}}$.

**7.1.7.** Fixing $i$, write

$$\widetilde{\tilde{F}}_i^{\mathfrak{p}} = \sum_{j=0}^{N} \widetilde{\tilde{a}}_j T_j.$$

Check the following condition:

$$\textit{For each } i \textit{ and each } j \leq N, \textit{ if } \widetilde{\tilde{a}}_j \neq 0 \textit{ then } |\widetilde{\tilde{a}}_j| \geq \epsilon.$$

If this *fails*, return, *False*.

**7.1.8.** We continue to work with the single $\epsilon$-ball $B$. Set $b$ equal to the number of points (0 or 1) in $X(Z)_n \cap B$. Choose an $r \in \mathbb{N}$ such that $\epsilon \geq p^{-r}$. Run the root-criterion algorithm $\mathcal{A}_{\mathrm{RC}}(b, N, r, h_i, \widetilde{\tilde{F}}_i^{\mathfrak{p}})$ for varying $i$.

[15] Evidently, this choice is arbitrary. In reality we would probably search up to a chosen height bound $B(n)$ which grows to $\infty$ as $n$ goes to $\infty$.

**7.1.9.** Repeat the steps of segements 7.1.6–7.1.9 for each $\epsilon$-ball $B$. If for each ball $B$ there exists an $i$ such that

$$\mathcal{A}_{\mathrm{RC}}(b, N, r, h_i, \widetilde{\tilde{F}_i^{\mathfrak{p}}}) = \mathit{True},$$

return *True*. Otherwise return *False*. This completes the construction of the algorithm.

**7.2.** ***Equation-solving theorem.*** We come to the main applications announced in the introduction.

**Theorem 7.2.1.** *Let $Z$ be an open integer scheme with fraction field $K$:*

(1) *Suppose the algorithm $\mathcal{A}_{\mathrm{ES}}(Z)$ halts. Then we have*

$$\mathcal{A}_{\mathrm{ES}}(Z) = X(Z).$$

(2) *Assume $K$ is totally real. Suppose Kim's conjecture* (Conjecture 2.1.4) *holds for $Z$ at level $n$. Suppose Zagier's conjecture* (Conjecture 2.2.5) *holds for $K$ and $n' \leq n$. Suppose Goncharov's conjecture* (Conjecture 2.2.7) *holds for $Z$ and $n' \leq n$. Suppose the period conjecture holds for the open subscheme $Z^o \subset Z$ constructed in segment 3.8 in half-weights $n' \leq n$. Suppose $K$ obeys the Hasse principle for finite cohomology* (segment 2.2.12) *in half-weights $2 \leq n' \leq n$. Then $\mathcal{A}_{\mathrm{ES}}(Z)$ halts.*

(3) *Assume $K = \mathbb{Q}$. Suppose Kim's conjecture holds for $Z$ at level $n$. Suppose Goncharov's conjecture holds for $Z$ and $n' \leq n$. Suppose the period conjecture holds for the open subscheme $Z^o \subset Z$ constructed in segment 3.8 in half-weights $n' \leq n$. Then $\mathcal{A}_{\mathrm{ES}}(Z)$ halts.*

*Proof.* Parts (1) and (2) are a direct application of Theorem 2.4.1. One point may require clarification: the role of segment 7.1.7. Let $\{F_i^{\mathfrak{p}}\}_i$ denote the generators of the Chabauty–Kim ideal close to $\{\tilde{F}_i^{\mathfrak{p}}\}_i$ whose existence is guaranteed by Theorem 2.4.1. Fixing an $\epsilon$-ball $B$ with representative $y \in B$ and an $i$, write

$$F_i^{\mathfrak{p}} = \sum_{j=1}^{\infty} a_j T^j \quad \text{and} \quad \tilde{F}_i^{\mathfrak{p}} = \sum_{j=0}^{\infty} \tilde{a}_j T^j$$

for the power series expansions about $y$. Then after an admissible change in $\epsilon$ (depending on $N$), we have

$$|a_j - \tilde{a}_j| < \epsilon$$

for all $j \leq N$. By construction, we have

$$|\tilde{a}_j - \widetilde{\tilde{a}_j}| < \epsilon,$$

hence

$$|a_j - \widetilde{\tilde{a}_j}| < \epsilon.$$

For part (1) of the theorem, suppose that for given $B$, $i$ and $j$, we find that $|\widetilde{\tilde{a}_j}| \geq \epsilon$. Then by the nonarchimedian triangle inequality, we have

$$|\widetilde{\tilde{a}_j}| = |a_j|.$$

This means that those valuations whose precise determination is needed for the root criterion algorithm $\mathcal{A}_{\mathrm{RC}}$, will indeed be precise. For part (2), we need only note that for $\epsilon$ sufficiently small depending on $N$, we will indeed have for each $\epsilon$-ball $B$, each $i$, and each $j \leq N$, either $\widetilde{\widetilde{a}}_j = 0$ or $|\widetilde{\widetilde{a}}_j| \geq \epsilon$.

We turn to part (3). The period conjecture implies in particular that the $p$-adic zeta values $\zeta^p(n')$ for $n' \in [3, n]$ odd are nonzero. In turn, this implies that the unipotent zeta values

$$\zeta^U(n') \in \mathrm{Ext}^1(\mathbb{Q}(0), \mathbb{Q}(n'))$$

are nonzero. Since these extension groups have dimension 1 while the remaining extension groups have dimension 0, this would imply Zagier's conjecture in this case. (In fact, since the extension groups for $n \geq 2$ don't depend on $Z$, we're free to choose any prime $p$. Choosing a regular one, where the nonvanishing is known, we obtain one possible proof of Zagier's conjecture for this case.)

We claim that the period conjecture also implies the Hasse principle. Indeed, the Hasse principle in this case merely says that a linear map from a vector space of dimension $\leq 1$ is injective. For this, it's enough to show that the map is nonzero. However, the composite map

$$\mathbb{Q}_p \otimes \mathrm{Ext}^1_{\mathbb{Z}}(\mathbb{Q}(0), \mathbb{Q}(n')) \to \mathrm{Ext}^1_{\mathbb{Z}_p}(\mathbb{Q}_p(0), \mathbb{Q}_p(n')) \to \mathbb{Q}_p$$

sends $\zeta^U(n')$ to $\zeta^p(n')$ which, as we've already noted, is nonzero if the period conjecture holds. □

## 8. Beyond totally real fields

**8.1.** We first explain the inadequacy of the methods developed above for dealing with fields which are not totally real.

**Proposition.** *Let $Z$ be an open integer scheme, $n$ a natural number, $\mathfrak{p} \in Z$ a totally split prime, and $X(\mathcal{O}_{\mathfrak{p}})_n$ the associated polylogarithmic Chabauty–Kim locus. Suppose $Z$ is not totally real, and assume the period conjecture holds. Then $X(\mathcal{O}_{\mathfrak{p}})_n = X(\mathcal{O}_{\mathfrak{p}})$.*

*Proof.* In this case, each motivic Ext group $\mathrm{Ext}^1_Z(\mathbb{Q}(0), \mathbb{Q}(n))$ is nonzero, and by the period conjecture, each (abelian) syntomic realization map

$$\mathrm{Ext}^1_Z(\mathbb{Q}(0), \mathbb{Q}(n)) \to \mathrm{Ext}^1_{\mathcal{O}_{\mathfrak{p}}}(\mathbb{Q}_p(0), \mathbb{Q}_p(n))$$

is at least nonzero. Since the motivic $\mathrm{Ext}^2$-groups are nevertheless zero, the Selmer scheme

$$H^1(G(Z), U(X)^{\mathrm{PL}}_{\geq -n})$$

is an $\mathrm{Ext}^1_Z(\mathbb{Q}(0), \mathbb{Q}(n))$-torsor over $H^1(G(Z), U(X)^{\mathrm{PL}}_{\geq -(n-1)})$. The realization map

$$\mathfrak{R}_{\mathfrak{p}} : H^1(G(Z), U(X)^{\mathrm{PL}}_{\geq -n}) \to H^1(G(\mathcal{O}_{\mathfrak{p}}), U(X)^{\mathrm{PL}}_{\geq -n})$$

is compatible with torsor structures. It follows by induction that $\mathfrak{R}_{\mathfrak{p}}$ is surjective, which completes the proof of the proposition. □

**8.2.** Instead of considering the single realization map $\mathfrak{R}_{\mathfrak{p}}$ we may consider the product, which fits into a square similar to the one considered above

$$\begin{array}{ccc} X(Z) & \longrightarrow & \prod_{\mathfrak{p}\,|\,p} X(\mathcal{O}_{\mathfrak{p}}) \\ \downarrow & & \downarrow \alpha \\ H^1(G(Z), U(X)^{\mathrm{PL}}_{\geq -n}) & \xrightarrow[\mathfrak{R}_p]{} & \prod_{\mathfrak{p}\,|\,p} H^1(G(\mathcal{O}_{\mathfrak{p}}), U(X)^{\mathrm{PL}}_{\geq -n}). \end{array}$$

We define the *big polylogarithmic Chabauty–Kim locus* by

$$\left(\prod_{\mathfrak{p}\,|\,p} X(\mathcal{O}_{\mathfrak{p}})\right)_n := \alpha^{-1}(\mathrm{Im}\,\mathfrak{R}_p).$$

and we again symmetrize with respect to the $S_3$ action to obtain a *symmetrized big polylogarithmic Chabauty–Kim locus*

$$\left(\prod_{\mathfrak{p}\,|\,p} X(\mathcal{O}_{\mathfrak{p}})\right)^{S_3}_n.$$

We also write $\mathcal{K}_p^{\mathrm{Big}}(\mathfrak{n}^{\mathrm{PL}}_{\geq -n})$ and $\mathcal{K}_p^{\mathrm{Big}}(\mathfrak{n}^{\mathrm{PL}}_{\geq -n})S_3$ for the corresponding ideals of Coleman functions, the *big $p$-adic Chabauty–Kim ideal* and *symmetrized big $p$-adic Chabauty–Kim ideal*, respectively. We restrict ourselves as usual to the totally split case for simplicity.

**Conjecture 8.2.1** (convergence of polylogarithmic loci, general case. Joint with David Corwin). *Let $Z$ be an open integer scheme and $p$ a prime below $Z$, for which $Z$ is totally split. For each $n \in \mathbb{N}$ let $\left(\prod_{\mathfrak{p}\,|\,p} X(\mathcal{O}_{\mathfrak{p}})\right)^{S_3}_n$ denote the associated symmetrized big polylogarithmic Chabauty–Kim locus. Then for some $n$ we have*

$$X(Z) = \left(\prod_{\mathfrak{p}\,|\,p} X(\mathcal{O}_{\mathfrak{p}})\right)^{S_3}_n.$$

**8.3.** A straightforward modification of the algorithm $\mathcal{A}_{\mathrm{Loci}}$ yields an algorithm which produces approximate generators for the ideal of polylogarithmic functions defining the big polylogarithmic loci. Its output includes an algebra basis $\mathcal{B}$ of $A(Z^o)_{\leq} n$ for $Z^o$ an open subscheme of $Z$, as well as a family of elements $\tilde{F}_i$ of the polynomial ring

$$\mathbb{Q}[\mathcal{B}, \{\log t_{\mathfrak{p}}\}_{\mathfrak{p}}, \{\mathrm{Li}_i\, t_{\mathfrak{p}}\}_{i,\mathfrak{p}}].$$

We denote its output by $\mathcal{A}_{\mathrm{Big-Loci}}(Z, p, n, \epsilon)$. The result is a theorem which is analogous to the one stated above.

**Theorem 8.3.1.** *Let $Z$ be an open integer scheme, $p$ a prime over which $Z$ it totally split, $n$ a natural number, and $\epsilon \in p^{\mathbb{Z}}$:*

(1) *Suppose* $\mathcal{A}_{\mathrm{Big-Loci}}(Z, p, n, \epsilon)$ *halts. Then there are functions* $\{F_i^{\mathfrak{p}}\}$ *generating the big p-adic Chabauty–Kim ideal* $\mathcal{K}_p^{\mathrm{Big}}(\mathfrak{n}^{\mathrm{PL}}_{\geq -n})$ *associated to* $\mathfrak{n}^{\mathrm{PL}}_{\geq -n}$ *such that*

$$|\tilde{F}_i^{\mathfrak{p}} - F_i^{\mathfrak{p}}| < \epsilon$$

*for all* $i$.

(2) *Suppose Zagier's conjecture* (Conjecture 2.2.5) *holds for* $K$ *and* $n' \leq n$. *Suppose Goncharov exhaustion* (Conjecture 2.2.7) *holds for* $Z$ *and* $n' \leq n$. *Suppose the period conjecture holds for the open subscheme* $Z^o \subset Z$ *constructed in segment 3.8 in half-weights* $n' \leq n$. *Suppose* $K$ *obeys the Hasse principle for finite cohomology* (segment 2.2.12) *in half-weights* $2 \leq n' \leq n$. *Then the computation* $\mathcal{A}_{\mathrm{Loci}}(Z, p, n, \epsilon)$ *halts.*

## Appendix: A minor erratum

**A.1.** The article [Goncharov 2001] contains a minor error: if Lemma 3.7 of that article were true, our algorithm could be greatly simplified. However, Clément Dupont has pointed out the following simple counterexample: $(\log^U 2)\zeta^U(3)$ is ramified at 2. To see this, note that $A(\operatorname{Spec}\mathbb{Z})_4 = 0$, that $A(\operatorname{Spec}\mathbb{Q})$ is an integral domain, and that both $\log^U 2$ and $\zeta^U(3)$ are nonzero. However, since both of these elements are contained in the space of extensions, in the notation of that article, we have $\Delta'_{[4]}((\log^U 2)\zeta^U(3)) = 0$.

## Acknowledgements

I would like to thank Stefan Wewers for helpful conversations during the conference on multiple zeta values in Madrid in December of 2014. I would like to thank Minhyong Kim, Amnon Besser, Francis Brown, Francesc Fité, Go Yamashita for helpful conversations and email exchanges. I would like to thank Clément Dupont for long conversations during our time in Sarrians, and for pointing out a very helpful counterexample (see the appendix). I would like to thank Rodolfo Venerucci for conversations about finite cohomology. I would like to thank Jochen Heinloth and Giuseppe Ancona for help improving my presentation of the results. I would like to thank David Corwin for a careful reading and many helpful comments; moreover, in the course of our joint work [Corwin and Dan-Cohen 2018a], we discovered that Conjecture 2.1.4 was false as stated in a previous draft. Finally, I wish to thank the referees for their helpful comments and suggestions.

## References

[Baker and Wüstholz 2007] A. Baker and G. Wüstholz, *Logarithmic forms and Diophantine geometry*, New Math. Monogr. **9**, Cambridge Univ. Press, 2007. MR Zbl

[Balakrishnan et al. 2018] J. S. Balakrishnan, I. Dan-Cohen, M. Kim, and S. Wewers, "A non-abelian conjecture of Tate–Shafarevich type for hyperbolic curves", *Math. Ann.* **372**:1-2 (2018), 369–428. MR Zbl

[Beilinson 1989] A. A. Beilinson, "Polylogarithm and cyclotomic elements", preprint, 1989.

[Besser 2002] A. Besser, "Coleman integration using the Tannakian formalism", *Math. Ann.* **322**:1 (2002), 19–48. MR Zbl

[Besser 2012] A. Besser, "Heidelberg lectures on Coleman integration", pp. 3–52 in *The arithmetic of fundamental groups* (Heidelberg, 2010), edited by J. Stix, Contrib. Math. Comput. Sci. **2**, Springer, 2012. MR Zbl

[Besser and de Jeu 2008] A. Besser and R. de Jeu, "$\mathrm{Li}^{(p)}$-service? An algorithm for computing $p$-adic polylogarithms", *Math. Comp.* **77**:262 (2008), 1105–1134. MR Zbl

[Bloch 2000] S. J. Bloch, *Higher regulators, algebraic K-theory, and zeta functions of elliptic curves*, CRM Monogr. Series **11**, Amer. Math. Soc., Providence, RI, 2000. MR Zbl

[Bloch and Kato 1990] S. Bloch and K. Kato, "$L$-functions and Tamagawa numbers of motives", pp. 333–400 in *The Grothendieck Festschrift, I*, edited by P. Cartier et al., Progr. Math. **86**, Birkhäuser, Boston, 1990. MR Zbl

[Borel 1953] A. Borel, "Sur la cohomologie des espaces fibrés principaux et des espaces homogènes de groupes de Lie compacts", *Ann. of Math.* (2) **57** (1953), 115–207. MR Zbl

[Borel 1977] A. Borel, "Cohomologie de $\mathrm{SL}_n$ et valeurs de fonctions zeta aux points entiers", *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **4**:4 (1977), 613–636. MR Zbl

[Brown 2012] F. C. S. Brown, "On the decomposition of motivic multiple zeta values", pp. 31–58 in *Galois–Teichmüller theory and arithmetic geometry*, edited by H. Nakamura et al., Adv. Stud. Pure Math. **63**, Math. Soc. Japan, Tokyo, 2012. MR Zbl

[Brown 2013] F. Brown, "Single-valued periods and multiple zeta values", preprint, 2013. arXiv

[Brown 2017] F. Brown, "Integral points on curves, the unit equation, and motivic periods", preprint, 2017. arXiv

[Chatzistamatiou 2017] A. Chatzistamatiou, "On integrality of $p$-adic iterated integrals", *J. Algebra* **474** (2017), 240–270. MR Zbl

[Chatzistamatiou and Ünver 2013] A. Chatzistamatiou and S. Ünver, "On $p$-adic periods for mixed Tate motives over a number field", *Math. Res. Lett.* **20**:5 (2013), 825–844. MR Zbl

[Corwin and Dan-Cohen 2018a] D. Corwin and I. Dan-Cohen, "The polylog quotient and the Goncharov quotient in computational Chabauty–Kim theory, I", preprint, 2018. To appear in *International Journal of Number Theory*. arXiv

[Corwin and Dan-Cohen 2018b] D. Corwin and I. Dan-Cohen, "The polylog quotient and the Goncharov quotient in computational Chabauty–Kim theory, II", preprint, 2018. To appear in *Transactions of the American Mathematical Society*. arXiv

[Dan-Cohen and Chatzistamatiou 2014] I. Dan-Cohen and A. Chatzistamatiou, "Computation of $p$-adic iterated integrals", unpublished, 2014.

[Dan-Cohen and Wewers 2015] I. Dan-Cohen and S. Wewers, "Explicit Chabauty–Kim theory for the thrice punctured line in depth 2", *Proc. Lond. Math. Soc.* (3) **110**:1 (2015), 133–171. MR Zbl

[Dan-Cohen and Wewers 2016] I. Dan-Cohen and S. Wewers, "Mixed Tate motives and the unit equation", *Int. Math. Res. Not.* **2016**:17 (2016), 5291–5354. MR Zbl

[Deligne 1989] P. Deligne, "Le groupe fondamental de la droite projective moins trois points", pp. 79–297 in *Galois groups over* $\mathbb{Q}$ (Berkeley, 1987), edited by Y. Ihara et al., Math. Sci. Res. Inst. Publ. **16**, Springer, 1989. MR Zbl

[Deligne 2010] P. Deligne, "Le groupe fondamental unipotent motivique de $\mathbb{G}_m - \mu_N$, pour $N = 2, 3, 4, 6$ ou 8", *Publ. Math. Inst. Hautes Études Sci.* **112** (2010), 101–141. MR Zbl

[Deligne and Goncharov 2005] P. Deligne and A. B. Goncharov, "Groupes fondamentaux motiviques de Tate mixte", *Ann. Sci. École Norm. Sup.* (4) **38**:1 (2005), 1–56. MR Zbl

[Evertse and Győry 2015] J.-H. Evertse and K. Győry, *Unit equations in Diophantine number theory*, Cambridge Stud. Adv. Math. **146**, Cambridge Univ. Press, 2015. MR Zbl

[Furusho 2004] H. Furusho, "$p$-adic multiple zeta values, I: $p$-adic multiple polylogarithms and the $p$-adic KZ equation", *Invent. Math.* **155**:2 (2004), 253–286. MR Zbl

[Furusho 2007] H. Furusho, "$p$-adic multiple zeta values, II: Tannakian interpretations", *Amer. J. Math.* **129**:4 (2007), 1105–1144. MR Zbl

[Goncharov 1994] A. B. Goncharov, "Polylogarithms and motivic Galois groups", pp. 43–96 in *Motives* (Seattle, WA, 1991)), edited by U. Jannsen et al., Proc. Sympos. Pure Math. **55**, Amer. Math. Soc., Providence, RI, 1994. MR

[Goncharov 1995] A. B. Goncharov, "Polylogarithms in arithmetic and geometry", pp. 374–387 in *Proc. Int. Congress of Math., I* (Zürich, 1994), edited by S. D. Chatterji, Birkhäuser, Basel, 1995. MR Zbl

[Goncharov 2001] A. B. Goncharov, "Multiple polylogarithms and mixed Tate motives", preprint, 2001. arXiv

[Goncharov 2005] A. B. Goncharov, "Galois symmetries of fundamental groupoids and noncommutative geometry", *Duke Math. J.* **128**:2 (2005), 209–284. MR Zbl

[Jannsen 1989] U. Jannsen, "On the $l$-adic cohomology of varieties over number fields and its Galois cohomology", pp. 315–360 in *Galois groups over* $\mathbb{Q}$ (Berkeley, 1987), edited by Y. Ihara et al., Math. Sci. Res. Inst. Publ. **16**, Springer, 1989. MR Zbl

[Jarossay 2016] D. Jarossay, "Pro-unipotent harmonic actions and a dynamical method for the computation of $p$-adic cyclotomic multiple zeta values", preprint, 2016. arXiv

[von Känel and Matschke 2016] R. von Känel and B. Matschke, "Solving $S$-unit, Mordell, Thue, Thue–Mahler and generalized Ramanujan–Nagell equations via Shimura–Taniyama conjecture", preprint, 2016. arXiv

[Kim 2005] M. Kim, "The motivic fundamental group of $\mathbb{P}^1 \setminus \{0, 1, \infty\}$ and the theorem of Siegel", *Invent. Math.* **161**:3 (2005), 629–656. MR Zbl

[Kim 2009] M. Kim, "The unipotent Albanese map and Selmer varieties for curves", *Publ. Res. Inst. Math. Sci.* **45**:1 (2009), 89–133. MR Zbl

[Kim 2012] M. Kim, "Tangential localization for Selmer varieties", *Duke Math. J.* **161**:2 (2012), 173–199. MR Zbl

[Perrin-Riou 1994] B. Perrin-Riou, "La fonction $L$ $p$-adique de Kubota–Leopoldt", pp. 65–93 in *Arithmetic geometry* (Tempe, AZ, 1993), edited by N. Childress and J. W. Jones, Contemp. Math. **174**, Amer. Math. Soc., Providence, RI, 1994. MR Zbl

[Soulé 1981] C. Soulé, "On higher $p$-adic regulators", pp. 372–401 in *Algebraic $K$-theory* (Evanston, IL, 1980), edited by E. M. Friedlander and M. R. Stein, Lecture Notes in Math. **854**, Springer, 1981. MR Zbl

[Suslin 1987] A. A. Suslin, "Algebraic $K$-theory of fields", pp. 222–244 in *Proc. Int. Congress of Math., I* (Berkeley, 1986), edited by A. M. Gleason, Amer. Math. Soc., Providence, RI, 1987. MR Zbl

[Ünver 2019] S. Ünver, "Cyclotomic $p$-adic multi-zeta values", *J. Pure Appl. Algebra* **223**:2 (2019), 489–503. MR Zbl

[de Weger 1989] B. M. M. de Weger, *Algorithms for Diophantine equations*, CWI Tract **65**, Stichting Math. Centrum, Amsterdam, 1989. MR Zbl

[Yamashita 2010] G. Yamashita, "Bounds for the dimensions of $p$-adic multiple $L$-value spaces", *Doc. Math.* extra volume (2010), 687–723. MR Zbl

[Zagier 1991] D. Zagier, "Polylogarithms, Dedekind zeta functions and the algebraic $K$-theory of fields", pp. 391–430 in *Arithmetic algebraic geometry* (Texel, Netherlands, 1989), edited by G. van der Geer et al., Progr. Math. **89**, Birkhäuser, Boston, 1991. MR Zbl

Communicated by Bjorn Poonen
Received 2018-10-10 Revised 2019-09-16 Accepted 2019-11-27

ishaidc@gmail.com *Department of Mathematics, Ben Gurion University, Be'er Sheva, Israel*

ALGEBRA AND NUMBER THEORY 14:5 (2020)
https://doi.org/10.2140/ant.2020.14.1239

# $p$-adic distribution of CM points and Hecke orbits I: Convergence towards the Gauss point

Sebastián Herrero, Ricardo Menares and Juan Rivera-Letelier

We study the asymptotic distribution of CM points on the moduli space of elliptic curves over $\mathbb{C}_p$, as the discriminant of the underlying endomorphism ring varies. In contrast with the complex case, we show that there is no uniform distribution. In this paper we characterize all the sequences of discriminants for which the corresponding CM points converge towards the Gauss point of the Berkovich affine line. We also give an analogous characterization for Hecke orbits. In the companion paper we characterize all the remaining limit measures of CM points and Hecke orbits.



## 1. Introduction

Given an algebraically closed field $\mathbb{K}$, denote by $Y(\mathbb{K})$ the moduli space of elliptic curves over $\mathbb{K}$. It is the space of all isomorphism classes of elliptic curves over $\mathbb{K}$, for isomorphisms defined over $\mathbb{K}$. For a class $E$ in $Y(\mathbb{K})$, the $j$-invariant $j(E)$ of $E$ is an element of $\mathbb{K}$ determining $E$ completely. The map $j\colon Y(\mathbb{K}) \to \mathbb{K}$ so defined is a bijection. See for example [Silverman 2009] and [Lang 1973] for background on elliptic curves.

If $\mathbb{K}$ is of characteristic 0, then the endomorphism ring of an elliptic curve defined over $\mathbb{K}$ is isomorphic to $\mathbb{Z}$ or to an order in a quadratic imaginary extension of $\mathbb{Q}$. In the latter case, the order only depends on the class $E$ in $Y(\mathbb{K})$ of the elliptic curve and $E$ is said to have *complex multiplication* or to be a *CM point*. In this paper, the *discriminant of a CM point* is the discriminant of the corresponding order.* Moreover, a

*MSC2010:* primary 11G15; secondary 11F32, 11S82.
*Keywords:* equidistribution, elliptic curves, Hecke correspondences.

*This notion of discriminant is not to be confused with the discriminant of a Weierstrass model of an elliptic curve [Silverman 2009, Chapter III, Section 1].

*discriminant* is the discriminant of an order in a quadratic imaginary extension of $\mathbb{Q}$. An integer $D$ is a discriminant if and only if $D < 0$ and $D \equiv 0, 1 \bmod 4$.

For every discriminant $D$, the set

$$\Lambda_D := \{E \in Y(\mathbb{K}) : \text{CM point of discriminant } D\} \tag{1-1}$$

is finite and nonempty. So, we can define the probability measure $\bar{\delta}_D$ on $Y(\mathbb{K})$, by

$$\bar{\delta}_D := \frac{1}{\#\Lambda_D} \sum_{E \in \Lambda_D} \delta_E,$$

where $\delta_x$ denotes the Dirac measure on $Y(\mathbb{K})$ at $x$.

Throughout the rest of this paper we fix a prime number $p$ and a completion $(\mathbb{C}_p, |\cdot|_p)$ of an algebraic closure of the field of $p$-adic numbers $\mathbb{Q}_p$. Our first goal is to study, for $\mathbb{K} = \mathbb{C}_p$, the asymptotic distribution of $\Lambda_D$ as the discriminant $D$ tends to $-\infty$. This is motivated by the following result in the case where $\mathbb{K}$ is the field of complex numbers $\mathbb{C}$. Recall that, if we consider the usual action of $\mathrm{SL}_2(\mathbb{Z})$ on the upper half-plane $\mathbb{H}$ by Möbius transformations, then $Y(\mathbb{C})$ can be naturally identified with the quotient space $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$. An appropriate multiple of the hyperbolic measure on $\mathbb{H}$ descends to a probability measure $\mu_{\mathrm{hyp}}$ on $Y(\mathbb{C})$.

**Theorem 1.** *For every continuous and bounded function $\varphi\colon Y(\mathbb{C}) \to \mathbb{R}$, we have*

$$\frac{1}{\#\Lambda_D} \sum_{E \in \Lambda_D} \varphi(E) \to \int \varphi \, \mathrm{d}\mu_{\mathrm{hyp}},$$

*as the discriminant $D$ tends to $-\infty$. Equivalently, we have the weak convergence of measures*

$$\bar{\delta}_D \to \mu_{\mathrm{hyp}},$$

*as the discriminant $D$ tends to $-\infty$.*

The asymptotic distribution of CM points on $Y(\mathbb{C})$ was part of a family of problems studied by Linnik; see [Linnik 1968] and also [Michel and Venkatesh 2006]. By applying a certain "ergodic method", Linnik proved the result above for sequences of discriminants satisfying some congruence restrictions. In a breakthrough, Duke [1988] removed the congruence restrictions assumed by Linnik and proved Theorem 1 for fundamental discriminants. Duke's proof uses the theory of nonholomorphic modular forms of half-integral weight and bounds for their Fourier coefficients, building on work of Iwaniec [1987]. Finally, Clozel and Ullmo [2004] obtained Theorem 1 for arbitrary discriminants, by studying the action of Hecke correspondences on CM points and combining Duke's result together with the uniform distribution of Hecke orbits.

**1A.** ***Convergence of CM points towards the Gauss point.*** Our first goal is to describe the asymptotic distribution of CM points for the ground field $\mathbb{K} = \mathbb{C}_p$. However, it is easy to find sequences of discriminants $(D_n)_{n=1}^{\infty}$ for which the sequence of measures $(\bar{\delta}_{D_n})_{n=1}^{\infty}$ on $Y(\mathbb{C}_p)$ has no accumulation measure. A natural solution to this issue is to consider $Y(\mathbb{C}_p)$ as a subspace of the Berkovich affine line

$\mathbb{A}^1_{\mathrm{Berk}}$ over $\mathbb{C}_p$, using the $j$-invariant to identify $Y(\mathbb{C}_p)$ with the subspace $\mathbb{C}_p$ of $\mathbb{A}^1_{\mathrm{Berk}}$. In fact, every sequence of measures $(\bar{\delta}_{D_n})_{n=1}^{\infty}$ as above accumulates on at least one probability measure with respect to the weak topology on the space of Borel measures on $\mathbb{A}^1_{\mathrm{Berk}}$. See Section 2D for a brief review of the space $\mathbb{A}^1_{\mathrm{Berk}}$ and the weak topology on the space of measures on $\mathbb{A}^1_{\mathrm{Berk}}$.

In contrast with Theorem 1, for $\mathbb{K} = \mathbb{C}_p$ the measures $\bar{\delta}_D$ on $\mathbb{A}^1_{\mathrm{Berk}}$ do not converge to a limit as the discriminant $D$ tends to $-\infty$. Our first main result is a characterization of all those sequences of discriminants $(D_n)_{n=1}^{\infty}$ tending to $-\infty$, such that the sequence of measures $(\bar{\delta}_{D_n})_{n=1}^{\infty}$ in $\mathbb{A}^1_{\mathrm{Berk}}$ converges to the Dirac measure at the "canonical" or "Gauss point" $x_{\mathrm{can}}$ of $\mathbb{A}^1_{\mathrm{Berk}}$. In the companion paper [Herrero et al. 2019] we show that in all the remaining cases the sequence $(\bar{\delta}_{D_n})_{n=1}^{\infty}$ accumulates on at least one probability measure supported on a compact subset of the supersingular locus of $Y(\mathbb{C}_p)$ and characterize all possible accumulation measures.

To state our first main result, we introduce some notation and terminology. Identify the residue field of $\mathbb{C}_p$ with an algebraic closure $\bar{\mathbb{F}}_p$ of the field with $p$ elements $\mathbb{F}_p$. Recall that the endomorphism ring of an elliptic curve over $\bar{\mathbb{F}}_p$ is isomorphic to an order in either a quadratic imaginary extension of $\mathbb{Q}$ or a quaternion algebra over $\mathbb{Q}$. In the former case the corresponding elliptic curve class is *ordinary* and it is *supersingular* in the latter.

Denote by $\mathcal{O}_p$ the ring of integers of $\mathbb{C}_p$ and by $\pi : \mathcal{O}_p \to \bar{\mathbb{F}}_p$ the reduction map. An elliptic curve class $E$ has *good reduction* if there is a representative Weierstrass equation with coefficients in $\mathcal{O}_p$ whose reduction is a smooth curve. Such reduction determines an elliptic curve defined over $\bar{\mathbb{F}}_p$, whose class $\tilde{E}$ only depends on $E$ and is the *reduction of* $E$. Moreover, $E$ has *ordinary* (resp. *supersingular*) *reduction* if $\tilde{E}$ is ordinary (resp. supersingular). An elliptic curve has good reduction precisely when $j(E)$ is in $\mathcal{O}_p$ and when this is not the case $E$ has *bad reduction*. The moduli space $Y(\mathbb{C}_p)$ is thus partitioned into three pairwise disjoint sets: The *bad*, *ordinary* and *supersingular reduction loci*, denoted by $Y_{\mathrm{bad}}(\mathbb{C}_p)$, $Y_{\mathrm{ord}}(\mathbb{C}_p)$ and $Y_{\mathrm{sups}}(\mathbb{C}_p)$, respectively. Using $j : Y(\mathbb{C}_p) \to \mathbb{C}_p$ to identify $Y(\mathbb{C}_p)$ and $\mathbb{C}_p$, we thus have the partition

$$\mathcal{O}_p = Y_{\mathrm{ord}}(\mathbb{C}_p) \sqcup Y_{\mathrm{sups}}(\mathbb{C}_p).$$

Moreover, if we denote by $Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)$ the finite subset of $Y(\bar{\mathbb{F}}_p)$ of supersingular classes, then $Y_{\mathrm{sups}}(\mathbb{C}_p) = \pi^{-1}(Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p))$ is a finite union of residue discs of $\mathcal{O}_p$. Note that $Y_{\mathrm{ord}}(\mathbb{C}_p)$ is a union of infinitely many residue discs of $\mathcal{O}_p$.

Every CM point $E$ has good reduction and the reduction type only depends on the discriminant $D$ of $E$, as follows:

(i) If $p$ splits in $\mathbb{Q}(\sqrt{D})$, then $E$ has ordinary reduction.

(ii) If $p$ ramifies or is inert in $\mathbb{Q}(\sqrt{D})$, then $E$ has supersingular reduction.

See [Deuring 1941] or [Lang 1973, Chapter 13, Section 4, Theorem 12]. We call a discriminant $D$ *$p$-ordinary* in the first case and *$p$-supersingular* in the second. Moreover, we define

$$|D|_{p\text{-sups}} := \begin{cases} 0 & \text{if } D \text{ is } p\text{-ordinary;} \\ |D|_p & \text{if } D \text{ is } p\text{-supersingular.} \end{cases}$$

**Theorem A.** *Let $(D_n)_{n=1}^{\infty}$ be a sequence of discriminants tending to $-\infty$. Then we have the weak convergence of measures*

$$\bar{\delta}_{D_n} \to \delta_{x_{\mathrm{can}}} \text{ as } n \to \infty \quad \text{if and only if} \quad |D_n|_{p\text{-sups}} \to 0 \text{ as } n \to \infty.$$

For readers unfamiliar with the Berkovich affine line, we give a concrete formulation of the convergence of measures in Theorem A in terms of $\mathbb{C}_p$ only, see Lemma 2.3(ii) in Section 2D.

We obtain Theorem A as a direct consequence of quantitative estimates in the cases where all the discriminants in $(D_n)_{n=1}^{\infty}$ are $p$-ordinary (Theorem 3.5 in Section 3B) or $p$-supersingular (Theorem 4.1 in Section 4). Note that in the former case Theorem A asserts that $\bar{\delta}_{D_n} \to \delta_{x_{\mathrm{can}}}$ weakly as $n \to \infty$. The following stronger statement is a direct consequence of our quantitative estimate in this case.

**Corollary B** (ordinary CM points are isolated). *Every disc of radius strictly less than one contained in $Y_{\mathrm{ord}}(\mathbb{C}_p)$ contains at most a finite number of CM points. In particular, the set of CM points in $Y_{\mathrm{ord}}(\mathbb{C}_p)$ is discrete.*

Corollary B seems to be well-known by the experts in the field, although we have not found this result explicitly stated in the literature. See Section 1C for comments and references.

**1B.** ***Convergence of Hecke orbits towards the Gauss point.*** To state our next main result, we first introduce Hecke correspondences. See Section 2B for background.

Given an algebraically closed field $\mathbb{K}$ of characteristic 0, a *divisor on* $Y(\mathbb{K})$ is an element of

$$\mathrm{Div}(Y(\mathbb{K})) := \bigoplus_{E \in Y(\mathbb{K})} \mathbb{Z}E,$$

the free abelian group spanned by the points of $Y(\mathbb{K})$. The *degree* and *support* of a divisor $\mathfrak{D} = \sum_{E \in Y(\mathbb{K})} n_E E$ in $\mathrm{Div}(Y(\mathbb{K}))$ are defined by

$$\deg(\mathfrak{D}) := \sum_{E \in Y(\mathbb{K})} n_E \quad \text{and} \quad \mathrm{supp}(\mathfrak{D}) := \{E \in Y(\mathbb{K}) : n_E \neq 0\},$$

respectively. If in addition $\deg(\mathfrak{D}) \geq 1$ and for every $E$ in $Y(\mathbb{K})$ we have $n_E \geq 0$, then

$$\bar{\delta}_{\mathfrak{D}} := \frac{1}{\deg(\mathfrak{D})} \sum_{E \in Y(\mathbb{K})} n_E \delta_E$$

is a probability measure on $Y(\mathbb{K})$.

For $n$ in $\mathbb{N} := \{1, 2, \ldots\}$ the $n$-th *Hecke correspondence* is the linear map

$$T_n \colon \mathrm{Div}(Y(\mathbb{K})) \to \mathrm{Div}(Y(\mathbb{K}))$$

defined for $E$ in $Y(\mathbb{K})$, by

$$T_n(E) := \sum_{C \leq E \text{ of order } n} E/C,$$

where the sum runs over all subgroups $C$ of $E$ of order $n$. Note that $\operatorname{supp}(T_n(E))$ is the set of all $E'$ in $Y(\mathbb{K})$ for which there is an isogeny $E \to E'$ of degree $n$. Moreover,

$$\deg(T_n(E)) = \sum_{d \mid n, d>0} d \geq n,$$

so $\deg(T_n(E)) \to \infty$ as $n \to \infty$.

In the case $\mathbb{K} = \mathbb{C}_p$, it is easy to see that for each $E$ in $Y_{\mathrm{bad}}(\mathbb{C}_p)$ (resp. $Y_{\mathrm{ord}}(\mathbb{C}_p)$, $Y_{\mathrm{sups}}(\mathbb{C}_p)$), we have that for every $n$ in $\mathbb{N}$ the divisor $T_n(E)$ is supported on $Y_{\mathrm{bad}}(\mathbb{C}_p)$ (resp. $Y_{\mathrm{ord}}(\mathbb{C}_p)$, $Y_{\mathrm{sups}}(\mathbb{C}_p)$).

**Theorem C.** *For every $E$ in $Y_{\mathrm{bad}}(\mathbb{C}_p) \cup Y_{\mathrm{ord}}(\mathbb{C}_p)$, we have the weak convergence of measures*

$$\bar{\delta}_{T_n(E)} \to \delta_{x_{\mathrm{can}}} \quad \text{as } n \to \infty.$$

*Moreover, for $E$ in $Y_{\mathrm{sups}}(\mathbb{C}_p)$ and a sequence $(n_j)_{j=1}^{\infty}$ in $\mathbb{N}$ tending to $\infty$, we have the weak convergence of measures*

$$\bar{\delta}_{T_{n_j}(E)} \to \delta_{x_{\mathrm{can}}} \text{ as } j \to \infty \quad \text{if and only if} \quad |n_j|_p \to 0 \text{ as } j \to \infty.$$

When restricted to the case where $E$ is in $Y_{\mathrm{bad}}(\mathbb{C}_p)$, the above theorem is [Richard 2018, Théorème 1.2].

To the best of our knowledge, Theorem C gives the first example where equidistribution of orbits fails for correspondences of degree bigger than one, see Section 2B for a description of Hecke correspondences as algebraic correspondences. In the complex case, pluripotential theory has been used successfully to prove equidistribution for correspondences satisfying a mild "nonmodularity" condition, see for example [Dinh et al. 2020].

The uniform distribution of Hecke orbits on $Y(\mathbb{C})$ is a well-known result from the spectral theory of automorphic forms; see [Clozel and Ullmo 2004, Théorème 2.1], and also [Clozel et al. 2001; Eskin and Oh 2006] for extensions and [Linnik and Skubenko 1964] for related work.

**Remark 1.1.** In [Clozel et al. 2001; Eskin and Oh 2006], the starting point is an algebraic group $G$ over $\mathbb{Q}$ and a congruence subgroup $\Gamma$ of $G(\mathbb{Q})$, and the ambient space is $X = \Gamma\backslash G(\mathbb{R})$. In this context, there is a natural notion of Hecke correspondences on $X$. The aforementioned works establish the uniform distribution of every orbit of such Hecke correspondences under general hypotheses. In particular, the $\mathbb{Q}$-structure of $G$ allows for $p$-adic variants of such results, see, e.g., [Clozel et al. 2001, Remark (1) in page 332]. In the particular case $G = \mathrm{SL}_2$ and $\Gamma = \mathrm{SL}_2(\mathbb{Z})$, there is a natural isomorphism $Y(\mathbb{C}) \simeq \mathrm{SL}_2(\mathbb{Z})\backslash \mathrm{SL}_2(\mathbb{R})/\mathrm{SO}_2(\mathbb{R})$ and the natural projection from $X$ to $Y(\mathbb{C})$ takes Hecke orbits as in [Clozel et al. 2001; Eskin and Oh 2006] to Hecke orbits on $Y(\mathbb{C})$ as defined in this paper. The uniform distribution of Hecke orbits on $Y(\mathbb{C})$ is thus a special case of [Clozel et al. 2001, Theorem 1.6], see also [Eskin and Oh 2006, Theorem 1.2]. However, this strategy breaks down for Hecke orbits on $Y(\mathbb{C}_p)$, because there is no analogous uniformization of $Y(\mathbb{C}_p)$ as a double quotient. Moreover, Theorem C shows that there is no uniform distribution of Hecke orbits on $Y(\mathbb{C}_p)$. Indeed, Theorem C and our results in the companion paper [Herrero et al. 2019] show that, in contrast with [Clozel et al. 2001; Clozel and

Ullmo 2004; Eskin and Oh 2006], the asymptotic distribution of $(T_{n_j}(E))_{j=1}^{\infty}$ on $Y(\mathbb{C}_p)$ depends on both the starting point $E$ and the sequence of integers $(n_j)_{j=1}^{\infty}$.

**1C.** ***Notes and references.*** After the first version of this paper was written, we learned about the related work of Goren and Kassaei [2017]. For a prime number $\ell$ different from $p$, Goren and Kassaei [2017] studied the dynamics of the Hecke correspondence $T_\ell$ acting on the moduli space of elliptic curves with a marked torsion point of exact order $N$ coprime to $p\ell$. So, on one hand [Goren and Kassaei 2017] is more general than this paper in that it considers modular curves with level structure. On the other hand, [loc. cit.] is more restrictive in that it only considers the dynamics of a single Hecke correspondence of prime index different from $p$, as opposed to the dynamics of the whole algebra of Hecke correspondences considered here. Note also that we use $\mathbb{C}_p$ as a ground field, which is natural to study equidistribution problems, whereas [loc. cit.] is restricted to algebraic extensions of $\mathbb{Q}_p$. In spite of the fact that both papers study the dynamics of similar maps, there is no significant intersection between the results of [loc. cit.] and those of this paper. See also [Herrero et al. 2019] for our additional results in the supersingular locus and the corresponding comparison with the results of [Goren and Kassaei 2017]. Finally, our results on the dynamics of the canonical branch $\boldsymbol{t}$ of $T_p$ (defined on $Y_{\mathrm{ord}}(\mathbb{C}_p)$ in Section 3A) on ordinary CM points show that this map gives rise to a "$(p+1)$-volcano" in the sense of [loc. cit., Section 2.1], see Remark 3.6.

Corollary B seems well-known among experts in the field, although we have not found this result explicitly stated in the literature. Even for higher-dimensional abelian varieties it can be deduced from the explicit characterization of the Serre–Tate local coordinates of CM points as torsion points of the multiplicative group, see, e.g., [de Jong and Noot 1991, Proposition 3.5]. Our approach makes no use of these local coordinates, and is based on rigid analytic properties of the canonical branch $\boldsymbol{t}$ of $T_p$. For CM elliptic curves with ordinary reduction, the connection between these two approaches is well-known, see, e.g., [Dwork 1969, Section 7d)].

Since every CM point of $Y(\mathbb{C}_p)$ is in the bounded set $\mathcal{O}_p$, Theorem A yields the following stronger statement: For every continuous function $\varphi\colon Y(\mathbb{C}_p) \to \mathbb{R}$ and every sequence of discriminants $(D_n)_{n=1}^{\infty}$ tending to $-\infty$ and satisfying $|D_n|_{p\text{-sups}} \to 0$ as $n \to \infty$, we have

$$\frac{1}{\#\deg(\Lambda_{D_n})} \sum_{E \in \Lambda_{D_n}} \varphi(E) \to \int \varphi \, \mathrm{d}\delta_{x_{\mathrm{can}}} \quad \text{as } n \to \infty.$$

Although our formulation of Theorem 1 seems stronger than the one in [Clozel and Ullmo 2004, Théorème 2.4], it is easy to see that it is equivalent, see for example [Bilu 1997, Lemma 2.2].

**1D.** ***Strategy and organization.*** We now explain the strategy of the proof of Theorems A and C and simultaneously describe the organization of the paper.

After some preliminaries in Section 2, we proceed to the proof of Theorem A in Sections 3 and 4. Theorem A is a direct consequence of stronger quantitative estimates in two separate cases: The case where all the discriminants in $(D_n)_{n=1}^{\infty}$ are $p$-ordinary and the case where they are all $p$-supersingular.

The $p$-ordinary case is treated in Section 3. There are two main ingredients, both of which are related to the "canonical branch $\boldsymbol{t}$" of $T_p$ that is defined in terms of the "canonical subgroup" in Section 3A; see also Appendix B. The first main tool is a simple formula, for every integer $m \geq 1$, of $T_{p^m}$ on $Y_{\mathrm{ord}}(\mathbb{C}_p)$ in terms of $\boldsymbol{t}$ (Proposition 3.4 in Section 3A). To establish this formula we use results of Tate and Deligne to show that $\boldsymbol{t}$ is rigid analytic. The second main tool is the interpretation of $p$-ordinary CM points as preperiodic points of $\boldsymbol{t}$ on $Y_{\mathrm{ord}}(\mathbb{C}_p)$ (Theorem 3.5(i)), which is based on Deuring's work on the canonical subgroup. Our quantitative estimate in the $p$-ordinary case is stated as Theorem 3.5(ii) in Section 3B and its proof is given at the end of this section.

The $p$-supersingular case is technically more difficult. We use Katz–Lubin's extension of the theory of canonical subgroups to "not too supersingular" elliptic curves and "Katz' valuation". We recall these in Section 4A, where we also give an explicit formula relating Katz' valuation to the $j$-invariant (Proposition 4.3). We use Katz' valuation to give a concrete description of the action of Hecke correspondences on the supersingular locus in terms of a sequence of correspondences $(\tau_m)_{m=1}^{\infty}$ acting on the interval $[0, p/(p+1)]$ (Proposition 4.5 in Section 4B). To do this, we rely on results in [Katz 1973, Section 3] and, for $p = 2$ and 3, on certain congruences satisfied by certain Eisenstein series, see Proposition A.1 in Appendix A. Our quantitative estimate in the $p$-ordinary case is stated as Theorem 4.1 at the beginning of Section 4 and its proof is given at the end of this section.

In Appendix B we formulate some of our results on the canonical branch $\boldsymbol{t}$ of $T_p$, as a lift of the classical Eichler–Shimura congruence relation (Theorem B.1).

The proof of Theorem C splits in three complementary cases, according to the reduction type of $E$. In each case we obtain a stronger quantitative estimate. For the bad reduction case we use Tate's uniformization theory (Proposition 5.1 in Section 5A). Thanks to the multiplicative properties of Hecke correspondences (2-6), the ordinary reduction case (Proposition 5.2 in Section 5B) is reduced to two special cases: The asymptotic distribution of $(T_{p^m}(E))_{m=1}^{\infty}$ (Proposition 5.3) and, for a sequence $(n_j)_{j=1}^{\infty}$ of integers in $\mathbb{N}$ that are not divisible by $p$, the asymptotic distribution of $(T_{n_j}(E))_{j=1}^{\infty}$ (Proposition 5.4). The former case is obtained using the tools developed in Theorem 3.5 and the latter is reduced to the study of the action of Hecke correspondences on ordinary elliptic curves in $Y(\bar{\mathbb{F}}_p)$ and is elementary. Finally, the supersingular case (Proposition 5.6 in Section 5C) is obtained from the description of the action of Hecke correspondences on the supersingular locus in Section 4B and an explicit formula for the correspondences $(\tau_m)_{m=1}^{\infty}$ (Lemma 5.7).

## 2. Preliminaries

Recall that $\mathbb{N} = \{1, 2, \ldots\}$. Given $n$ in $\mathbb{N}$, denote by

$$d(n) := \sum_{d>0, d \mid n} 1 \quad \text{and} \quad \sigma_1(n) := \sum_{d>0, d \mid n} d$$

the number and the sum of the positive divisors of $n$, respectively. We use several times the inequality

$$\sigma_1(n) \geq n, \tag{2-1}$$

and the fact that for every $\varepsilon > 0$ we have

$$d(n) = o(n^{\varepsilon}); \tag{2-2}$$

see for example [Apostol 1976, page 296].

For a set $X$ and a subset $A$ of $X$, we use $\mathbf{1}_A \colon X \to \{0, 1\}$ to denote the indicator function of $A$.

For a topological space $X$, denote by $\delta_x$ the *Dirac mass on $X$ supported at $x$*. It is the Borel probability measure characterized by the property that for every Borel subset $Y$ of $X$ we have $\delta_x(Y) = 1$ if $x \in Y$ and $\delta_x(Y) = 0$ otherwise.

Normalize the norm $|\cdot|_p$ of $\mathbb{C}_p$ so that $|p|_p = 1/p$ and denote by $\mathrm{ord}_p \colon \mathbb{C}_p \to \mathbb{R} \cup \{+\infty\}$ the valuation defined by $\mathrm{ord}_p(0) = +\infty$ and for $z$ in $\mathbb{C}_p^{\times}$ by $\mathrm{ord}_p(z) = -\log|z|_p/\log p$. Denote by $\mathcal{M}_p$ the maximal ideal of $\mathcal{O}_p$ and recall that we identify $\mathcal{O}_p/\mathcal{M}_p$ with $\bar{\mathbb{F}}_p$ and that $\pi \colon \mathcal{O}_p \to \bar{\mathbb{F}}_p$ denotes the reduction morphism. For $\zeta$ in $\bar{\mathbb{F}}_p$, denote by $\boldsymbol{D}(\zeta) := \pi^{-1}(\zeta)$ the residue disc corresponding to $\zeta$.

**2A.** ***Divisors.*** A *divisor* on a set $X$[†] is a formal finite sum $\sum_{x \in X} n_x x$ in $\bigoplus_{x \in X} \mathbb{Z}x$. In the special case where for some $x_0$ in $X$ we have $n_{x_0} = 1$ and $n_x = 0$ for every $x \neq x_0$, we use $[x_0]$ to denote this divisor. When there is no danger of confusion, sometimes we use $x_0$ to denote $[x_0]$.

Let $\mathfrak{D} = \sum_{x \in X} n_x[x]$ be a divisor on $X$. The *degree* and the *support* of $\mathfrak{D}$ are defined by

$$\deg(\mathfrak{D}) := \sum_{x \in X} n_x \quad \text{and} \quad \mathrm{supp}(\mathfrak{D}) := \{x \in X : n_x \neq 0\},$$

respectively. The divisor $\mathfrak{D}$ is *effective*, if for every $x$ in $X$ we have $n_x \geq 0$. For $A \subseteq X$, the *restriction of $\mathfrak{D}$ to $A$* is the divisor on $X$ defined by

$$\mathfrak{D}|_A := \sum_{x \in A} n_x[x].$$

For a set $X'$ and a map $f \colon X \to X'$, the *push-forward action of $f$ on divisors* $f_* \colon \mathrm{Div}(X) \to \mathrm{Div}(X')$ is the linear extension of the action of $f$ on points. In the particular case in which $X' = G$ is a commutative group, also define $f \colon \mathrm{Div}(X) \to G$ by

$$f(\mathfrak{D}) := \sum_{x \in X} n_x f(x) \in G.$$

If $X$ is a topological space and $\mathfrak{D}$ is an effective divisor satisfying $\deg(\mathfrak{D}) \geq 1$, then $\bar{\delta}_{\mathfrak{D}} := \frac{1}{\deg(\mathfrak{D})} \sum_{x \in X} n_x \delta_x$ is a Borel measure on $X$. Note that in the case $G = \mathbb{R}$ and $f$ is measurable, we have

$$\int f \, \mathrm{d}\bar{\delta}_{\mathfrak{D}} = \frac{f(\mathfrak{D})}{\deg(\mathfrak{D})}.$$

Since we are identifying $Y(\mathbb{C}_p)$ with $\mathbb{C}_p$ via $j$, we identify divisors on $Y(\mathbb{C}_p)$ and on $\mathbb{C}_p$ accordingly.

[†]We only use this definition in the case $X$ is one of several types of one-dimensional objects. For such $X$, the notion of divisor introduced here can be seen as a natural extension of the usual notion of Weil divisor.

**2B.** ***Hecke correspondences.*** In this section we recall the construction and main properties of the Hecke correspondences. For details we refer the reader to [Shimura 1971, Sections 7.2 and 7.3] for the general theory, or to the survey [Diamond and Im 1995, Part II].

Let $\mathbb{K}$ be an algebraically closed field of characteristic 0. First, note that for every integer $n \geq 1$ and divisor $\mathfrak{D}$ in $\mathrm{Div}(Y(\mathbb{K}))$, we have

$$\deg(T_n(\mathfrak{D})) = \sigma_1(n)\deg(\mathfrak{D}).$$

Moreover, for $n = 1$ the correspondence $T_1$ is by definition the identity on $\mathrm{Div}(Y(\mathbb{K}))$.

We also consider the linear extension of Hecke correspondences to $\mathrm{Div}(Y(\mathbb{K})) \otimes \mathbb{Q}$.

For an integer $N \geq 1$, denote by $Y_0(N)$ the *modular curve of level* $N$. It is a quasiprojective variety defined over $\mathbb{Q}$. The points of $Y_0(N)$ over $\mathbb{K}$ parametrize the moduli space of equivalence classes of pairs $(E, C)$, where $E$ is an elliptic curve over $\mathbb{K}$ and $C$ is a cyclic subgroup of $E$ of order $N$. Here, two such pairs $(E, C)$ and $(E', C')$ are equivalent if there exists an isomorphism $\phi\colon E \to E'$ over $\mathbb{K}$ taking $C$ to $C'$. In particular, when $N = 1$, for every algebraically closed field $\mathbb{K}$ we can parametrize $Y(\mathbb{K})$ by $Y_0(1)(\mathbb{K})$, and $Y_0(1)$ is isomorphic to the affine line $\mathbb{A}^1_{\mathbb{Q}}$.

For $N > 1$, denote by $\Phi_N(X, Y)$ the *modular polynomial of level* $N$, which is a symmetric polynomial in $\mathbb{Z}[X, Y]$ that is monic in both $X$ and $Y$, see, e.g., [Lang 1973, Chapter 5, Sections 2 and 3]. This polynomial is characterized by the equality

$$\Phi_N(j(E), Y) = \prod_{C \leq E \text{ cyclic of order } N} (Y - j(E/C)) \quad \text{for every } E \text{ in } Y(\mathbb{K}). \tag{2-3}$$

This implies that a birational model for $Y_0(N)$ is provided by the plane algebraic curve

$$\Phi_N(X, Y) = 0. \tag{2-4}$$

For each prime $q$, let $\alpha_q, \beta_q\colon Y_0(q) \to Y_0(1)$ be the rational maps over $\mathbb{Q}$ given in terms of moduli spaces by

$$\alpha_q(E, C) := E \quad \text{and} \quad \beta_q(E, C) := E/C.$$

In terms of the model (2-4) with $N = q$, the rational maps $\alpha_q$ and $\beta_q$ correspond to the projections on the $X$ and $Y$ coordinate, respectively. Denote by $(\alpha_q)_*$ and $(\beta_q)_*$ the push-forward action of $\alpha_q$ and $\beta_q$ on divisors, respectively, as in Section 2A. Denote also by $\alpha_q^*$ the pull-back action of $\alpha_q$ on divisors, defined at $x$ in $Y_0(1)(\mathbb{K})$ by

$$\alpha_q^*(x) := \sum_{\substack{y \in Y_0(q)(\mathbb{K}) \\ \alpha_q(y) = x}} \deg_{\alpha_q}(y)[y],$$

where $\deg_{\alpha_q}(y)$ is the local degree of $\alpha_q$ at $y$. This definition is extended by linearity to arbitrary divisors. The pull-back action $\beta_q^*$ of $\beta_q$ is defined in a similar way. Then the Hecke correspondence $T_q\colon \mathrm{Div}(Y(\mathbb{K})) \to \mathrm{Div}(Y(\mathbb{K}))$ is recovered as

$$T_q = (\alpha_q)_* \circ \beta_q^* = (\beta_q)_* \circ \alpha_q^*,$$

where the second equality follows from the first and from the symmetry of $T_q$.

For an arbitrary integer $n \geq 2$, the correspondence $T_n$ can be recovered from different $T_q$, for $q$ running over prime divisors of $n$, by using the identities

$$T_{q^r} = T_q \circ T_{q^{r-1}} - q \cdot T_{q^{r-2}} \quad \text{for } q \text{ prime and } r \geq 2; \tag{2-5}$$

$$T_\ell \circ T_m = T_{\ell m} \quad \text{for } \ell, m \geq 1 \text{ coprime.} \tag{2-6}$$

We conclude this section with the following lemma used in Sections 3A and 5B.

**Lemma 2.1.** *Let $n \geq 1$ be an integer. For $E$ in $Y(\mathbb{C}_p)$, the divisor $T_n(E)$ varies continuously with respect to $E$ in the following sense: For every commutative topological group $G$ and every continuous function $f : Y(\mathbb{C}_p) \to G$, the function $T_n f : Y(\mathbb{C}_p) \to G$ given by*

$$T_n f(E) := f(T_n(E))$$

*is continuous. In particular, for every open and closed subset $A \subseteq Y(\mathbb{C}_p)$, the integer valued map*

$$E \mapsto \deg(T_n(E)|_A)$$

*is locally constant.*

*Proof.* We first treat the case where $n$ equals a prime number $q$. Let $P_0(X), \ldots, P_q(X)$ be the polynomials in $\mathbb{Z}[X]$ such that

$$\Phi_q(X, Y) = P_0(X) + P_1(X)Y + \cdots + P_q(X)Y^q + Y^{q+1}.$$

Let $(E_m)_{m=1}^\infty$ be a sequence and $E_0$ be a point in $Y(\mathbb{C}_p)$, such that $j(E_m) \to j(E_0)$ when $m$ tends to infinity. Then for every $k$ in $\{0, 1, \ldots, q\}$, we have $P_k(j(E_m)) \to P_k(j(E_0))$ when $m$ tends to infinity. It follows that the roots of the polynomial $\Phi_q(j(E_m), Y)$ converge to the roots of $\Phi_q(j(E_0), Y)$, in the following sense: For every $m$ in $\{0, 1, 2, \ldots\}$ we can find $z_{m,0}, \ldots, z_{m,q}$ in $\mathbb{C}_p$, so that

$$\Phi_q(j(E_m), Y) = \prod_{k=0}^{q} (Y - z_{m,k}),$$

and so that for every $k$ in $\{0, 1, \ldots, q\}$ we have $z_{m,k} \to z_{0,k}$ when $m$ tends to infinity, see for example [Brink 2006, Theorem 2]. For each $m$ in $\{0, 1, 2, \ldots\}$ and $k$ in $\{0, 1, \ldots, q\}$, let $E_{m,k}$ be the curve in $Y(\mathbb{C}_p)$ with $j(E_{m,k}) = z_{m,k}$. By the definition of $T_q$ and (2-3), we have for every $m \geq 0$

$$T_q(E_m) = \sum_{k=0}^{q} [E_{m,k}].$$

Since for every $k$ in $\{0, 1, \ldots, q\}$ we have $j(E_{m,k}) \to j(E_{0,k})$ when $m$ tends to infinity, we conclude that for every continuous function $f : Y(\mathbb{C}_p) \to G$ we have

$$T_q f(E_m) = \sum_{k=0}^{q} f(E_{k,m}) \to \sum_{k=0}^{q} f(E_{k,0}) = T_q f(E_0).$$

This proves that $T_q f$ is continuous.

We now treat the general case by using multiplicative induction, the relations (2-5) and (2-6), and the fact that for every pair of linear maps $L, \tilde{L}\colon \mathrm{Div}(Y(\mathbb{C}_p)) \to \mathrm{Div}(Y(\mathbb{C}_p))$, every pair of integers $m, \tilde{m}$, and every function $F\colon Y(\mathbb{C}_p) \to G$, one has

$$(L \circ \tilde{L})(F) = \tilde{L}(L(F)) \quad \text{and} \quad (mL + \tilde{m}\tilde{L})(F) = mL(F) + \tilde{m}\tilde{L}(F). \tag{2-7}$$

Denote by $I$ the set of those integers $n \geq 1$ such that for every continuous function $f\colon Y(\mathbb{C}_p) \to G$, the function $T_n(f)$ is also continuous. Clearly $I$ contains 1, since for every function $f$ we have $T_1(f) = f$. By the proof given above, $I$ contains all prime numbers. Let $n \geq 1$ be a given integer having each divisor in $I$, and let $q$ be a prime number. Let $s \geq 0$ and $n_0 \geq 1$ be the integers such that $n = q^s n_0$, and such that $q$ does not divide $n_0$. Then by the relations (2-5) and (2-6), and by (2-7), we have

$$T_{qn}(f) = T_{q^{s+1}n_0}(f) = T_{n_0}(T_{q^{s+1}}(f)),$$

and for $s \geq 1$

$$T_{q^{s+1}}(f) = T_{q^s}(T_q(f)) - qT_{q^{s-1}}(f).$$

Since $n_0$, $q$, $q^s$, and $q^{s-1}$ if $s \geq 1$, are all in $I$, we conclude that $T_{qn}(f)$ is continuous, and that $qn$ is in $I$. This completes the proof of the multiplicative induction step, and of the first part of the lemma.

The second part of the lemma is an easy consequence of the first. Indeed, let $A \subseteq Y(\mathbb{C}_p)$ be an open and closed subset. Then the function $\mathbf{1}_A$ is continuous and the first part implies that

$$E \mapsto T_n\mathbf{1}_A(E) = \mathbf{1}_A(T_n(E)) = \deg(T_n(E)|_A)$$

is also continuous. But $T_n\mathbf{1}_A$ has integer values, hence it must be locally constant. This completes the proof of the lemma. □

**2C.** ***Hecke orbits of CM points and an estimate on class numbers.*** In this section we first recall a special case of a formula of Zhang describing the effect of Hecke correspondences on CM points (Lemma 2.2), which is used in Sections 3, 4 and 5B. To do this, and for the rest of the paper, for every discriminant $D$ we consider $\Lambda_D$ as a divisor. We also use Siegel's classical lower bound on class numbers of quadratic imaginary extensions of $\mathbb{Q}$, to give the following estimate used in the proof of Theorem A: For every $\varepsilon > 0$ there is a constant $C > 0$ such that for every negative discriminant $D$, we have

$$h(D) := \deg(\Lambda_D) \geq C|D|^{1/2-\varepsilon}. \tag{2-8}$$

In this section we follow [Clozel and Ullmo 2004, Section 2.3], adding some details for the benefit of the reader.

We use $d$ to denote a negative fundamental discriminant. For each discriminant $D$ there is a unique negative fundamental discriminant $d$ and integer $f \geq 1$ such that $D = df^2$. These are the *fundamental discriminant* and *conductor of* $D$, respectively. We denote by $\mathcal{O}_{d,f}$ the unique order of discriminant $D$ in the quadratic imaginary extension $\mathbb{Q}(\sqrt{d})$ of $\mathbb{Q}$ and put

$$w_{d,f} := \#(\mathcal{O}_{d,f}^\times/\mathbb{Z}^\times) = (\#\mathcal{O}_{d,f}^\times)/2.$$

The integer $f$ is the index of $\mathcal{O}_{d,f}$ inside the ring of integers of $\mathbb{Q}(\sqrt{d})$. Note that $w_{-3,1}=3$, $w_{-4,1}=2$, and that in all the remaining cases $w_{d,f}=1$.

Recall that the *Dirichlet convolution* of two functions $g, \tilde{g}\colon \mathbb{N}\to\mathbb{C}$, is defined by

$$(g*\tilde{g})(n) := \sum_{d\in\mathbb{N}, d\mid n} g(d)\tilde{g}\left(\frac{n}{d}\right).$$

Given a fundamental discriminant $d$, denote by $R_d\colon \mathbb{N}\to\mathbb{N}\cup\{0\}$ the function that to each $n$ in $\mathbb{N}$ assigns the number of integral ideals of norm $n$ in the ring of integers of $\mathbb{Q}(\sqrt{d})$. Moreover, denote by $R_d^{-1}$ the inverse of $R_d$ with respect to the Dirichlet convolution.

**Lemma 2.2.** *For every fundamental discriminant $d<0$ and any pair of coprime integers $f\geq 1$ and $\tilde{f}\geq 1$, we have the relations*

$$T_f\left(\frac{\Lambda_{d\tilde{f}^2}}{w_{d,\tilde{f}}}\right) = \sum_{f_0\in\mathbb{N}, f_0\mid f} R_d\left(\frac{f}{f_0}\right)\frac{\Lambda_{d(f_0\tilde{f})^2}}{w_{d,f_0\tilde{f}}}; \tag{2-9}$$

$$\frac{\Lambda_{d(f\tilde{f})^2}}{w_{d,f\tilde{f}}} = \sum_{f_0\in\mathbb{N}, f_0\mid f} R_d^{-1}\left(\frac{f}{f_0}\right)T_{f_0}\left(\frac{\Lambda_{d\tilde{f}^2}}{w_{d,\tilde{f}}}\right). \tag{2-10}$$

*If in addition $f$ is not divisible by $p$, then we have*

$$\Lambda_{d(pf)^2} = \begin{cases} T_p\left(\frac{\Lambda_{df^2}}{w_{d,f}}\right) - 2\frac{\Lambda_{df^2}}{w_{d,f}} & \text{if } p \text{ splits in } \mathbb{Q}(\sqrt{d});\\ T_p\left(\frac{\Lambda_{df^2}}{w_{d,f}}\right) - \frac{\Lambda_{df^2}}{w_{d,f}} & \text{if } p \text{ ramifies in } \mathbb{Q}(\sqrt{d});\\ T_p\left(\frac{\Lambda_{df^2}}{w_{d,f}}\right) & \text{if } p \text{ is inert in } \mathbb{Q}(\sqrt{d}), \end{cases} \tag{2-11}$$

*and for every integer $m\geq 2$ we have*

$$\Lambda_{d(p^m f)^2} = \begin{cases} T_{p^m}\left(\frac{\Lambda_{df^2}}{w_{d,f}}\right) - 2T_{p^{m-1}}\left(\frac{\Lambda_{df^2}}{w_{d,f}}\right) + T_{p^{m-2}}\left(\frac{\Lambda_{df^2}}{w_{d,f}}\right) & \text{if } p \text{ splits in } \mathbb{Q}(\sqrt{d});\\ T_{p^m}\left(\frac{\Lambda_{df^2}}{w_{d,f}}\right) - T_{p^{m-1}}\left(\frac{\Lambda_{df^2}}{w_{d,f}}\right) & \text{if } p \text{ ramifies in } \mathbb{Q}(\sqrt{d});\\ T_{p^m}\left(\frac{\Lambda_{df^2}}{w_{d,f}}\right) - T_{p^{m-2}}\left(\frac{\Lambda_{df^2}}{w_{d,f}}\right) & \text{if } p \text{ is inert in } \mathbb{Q}(\sqrt{d}). \end{cases} \tag{2-12}$$

To prove this lemma, we first record the following identity, which is also used in the proof (2-8) below and of Lemma 5.5 in Section 5B. Let $\psi_d$ be the quadratic character associated to $K=\mathbb{Q}(\sqrt{d})$, which is given by the Kronecker symbol $\left(\frac{d}{\cdot}\right)$, and denote by $\mathbf{1}\colon \mathbb{N}\to\mathbb{C}$ the constant function equal to 1. Then we have the equality of functions

$$R_d = \psi_d * \mathbf{1}. \tag{2-13}$$

In fact, if we denote by $\zeta(s)$ the Riemann zeta function, by $\zeta_K(s)$ the Dedekind zeta function associated to $K$, and by $L(\psi_d, s)$ the Dedekind $L$-function associated to $\psi_d$, then the formula above is equivalent to the factorization $\zeta_K(s)=\zeta(s)L(\psi_d,s)$, whose proof can be found for example in [Cohen 2007, Proposition 10.5.5 on page 219], or [Lang 1994, Chapter XII, Section 1, Theorem 1].

*Proof of Lemma 2.2.* From the Möbius inversion formula we deduce that (2-9) and (2-10) are equivalent. Hence, it is enough to prove (2-10). We have the following formula of Zhang

$$T_f\left(\frac{\Lambda_d}{w_{d,1}}\right) = \sum_{f_0\in\mathbb{N},\, f_0 \mid f} R_d\left(\frac{f}{f_0}\right)\frac{\Lambda_{df_0^2}}{w_{d,f_0}}, \tag{2-14}$$

see for example [Clozel and Ullmo 2004, Lemme 2.6] or [Zhang 2001, Proposition 4.2.1]. Applying the Möbius inversion formula, one obtains

$$\frac{\Lambda_{df^2}}{w_{d,f}} = \sum_{f_0\in\mathbb{N},\, f_0 \mid f} R_d^{-1}\left(\frac{f}{f_0}\right) T_{f_0}\left(\frac{\Lambda_d}{w_{d,1}}\right). \tag{2-15}$$

On the other hand, note that if $f$ and $\tilde{f}$ in $\mathbb{N}$ are coprime, then by (2-6) and (2-15), we obtain (2-10).

Finally, (2-11) and (2-12) are a direct consequence of (2-9), (2-13) and the fact that $\psi_d(p) = 1$ (resp. 0, $-1$) if $p$ splits (resp. ramifies, is inert) in $\mathbb{Q}(\sqrt{d})$. □

To prove (2-8), recall from the theory of complex multiplication that for a fundamental discriminant $d$ the number $h(d)$ equals the class number of the quadratic extension $\mathbb{Q}(\sqrt{d})$ of $\mathbb{Q}$, see for example [Cox 2013, Corollary 10.20]. A celebrated result by Siegel states that for every $\varepsilon > 0$ there exists a constant $C > 0$ such that for every fundamental discriminant $d < 0$ we have

$$h(d) \geq C|d|^{1/2-\varepsilon}, \tag{2-16}$$

see for example [Siegel 1935], or [Lang 1994, Chapter XVI, Section 4, Theorem 4]. On the other hand, by [Lang 1973, Chapter 8, Section 1, Theorem 7] for every integer $f \geq 2$ we have

$$h(df^2) = \frac{w_{d,f}}{w_d} h(d) f \prod_{q \mid f,\ \text{prime}} \left(\frac{q - \psi_d(q)}{q}\right). \tag{2-17}$$

Given $\varepsilon > 0$, there are $C'$ in $]0, 1[$ and $N$ in $\mathbb{N}$ such that $(q-1)/q \geq q^{-\varepsilon}$ for every $q > N$ and $(q-1)/q \geq C' q^{-\varepsilon}$ for every $2 \leq q \leq N$. Hence, for every integer $f \geq 2$ we have

$$\prod_{q \mid f,\ \text{prime}} \left(\frac{q - \psi_d(q)}{q}\right) \geq \prod_{q \mid f,\ \text{prime}} \left(\frac{q-1}{q}\right) \geq (C')^N \prod_{q \mid f,\ \text{prime}} q^{-\varepsilon} \geq (C')^N f^{-\varepsilon}.$$

Combined with (2-16) and (2-17), this completes the proof of (2-8).

**2D.** ***The Berkovich affine line over $\mathbb{C}_p$ and the Gauss point.*** We refer the reader to [Berkovich 1990] for the general theory of Berkovich spaces, and to [Baker and Rumely 2010, Chapter 1] for the special case of the Berkovich affine line over $\mathbb{C}_p$, which is the only Berkovich space of relevance in this paper.

The *Berkovich affine line over* $\mathbb{C}_p$, which we denote by $\mathbb{A}^1_{\mathrm{Berk}}$, is a topological space defined as follows: As a set, $\mathbb{A}^1_{\mathrm{Berk}}$ is the collection of all multiplicative seminorms on the polynomial ring $\mathbb{C}_p[X]$ that take values in $\mathbb{R}_0^+$ and that extend the $p$-adic norm $|\cdot|_p$ on $\mathbb{C}_p$. Hence, a point $x \in \mathbb{A}^1_{\mathrm{Berk}}$ is given by a map

$x\colon \mathbb{C}_p[X] \to \mathbb{R}_0^+$ satisfying for every $a$ in $\mathbb{C}_p$ and for all $f$ and $g$ in $\mathbb{C}_p[X]$,

$$x(a) = |a|_p, \quad x(f+g) \le x(f) + x(g) \quad \text{and} \quad x(fg) = x(f)x(g).$$

The topology of $\mathbb{A}^1_{\mathrm{Berk}}$ is the weakest topology such that for every $f \in \mathbb{C}_p[X]$, the function $\mathbb{A}^1_{\mathrm{Berk}} \to \mathbb{C}_p$ given by $x \mapsto x(f)$ is continuous. The topological space $\mathbb{A}^1_{\mathrm{Berk}}$ is Hausdorff, locally compact, metrizable and path-connected. It contains $\mathbb{C}_p$ as a dense subspace via the map $\iota\colon \mathbb{C}_p \to \mathbb{A}^1_{\mathrm{Berk}}$ given, for $z \in \mathbb{C}_p$ and $f \in \mathbb{C}_p[X]$, by $\iota(z)(f) := |f(z)|_p$. We identify divisors on $\mathbb{C}_p$ and on $\iota(\mathbb{C}_p)$ accordingly.

The *canonical point* or *Gauss point* $x_{\mathrm{can}}$ of $\mathbb{A}^1_{\mathrm{Berk}}$ is the Gauss norm

$$\sum_{n=0}^{N} a_n X^n \mapsto \sup\left\{ \left| \sum_{n=0}^{N} a_n z^n \right|_p : z \in \mathcal{O}_p \right\} = \max\{|a_n|_p : n \in \{0, \dots, N\}\}.$$

Given $a \in \mathbb{C}_p$ and $r > 0$, define

$$\begin{aligned} \boldsymbol{D}(a, r) &:= \{x \in \mathbb{C}_p : |x - a|_p < r\}; \\ \boldsymbol{D}^\infty(a, r) &:= \{x \in \mathbb{C}_p : |x - a|_p > r\}; \\ \mathcal{D}(a, r) &:= \{x \in \mathbb{A}^1_{\mathrm{Berk}} : x(X - a) < r\}; \\ \mathcal{D}^\infty(a, r) &:= \{x \in \mathbb{A}^1_{\mathrm{Berk}} : x(X - a) > r\}. \end{aligned}$$

A basis of neighborhoods of $x_{\mathrm{can}}$ in $\mathbb{A}^1_{\mathrm{Berk}}$ is given by the collection of sets

$$\mathcal{A}(A; R) := \mathcal{D}(0, R) \cap \bigcap_{a \in A} \mathcal{D}^\infty(a, R^{-1}), \tag{2-18}$$

where $R > 1$ and $A$ is a finite subset of $\mathcal{O}_p$.

We conclude this section with the following result. Recall that a sequence of Borel probability measures $(\mu_n)_{n\in\mathbb{N}}$ on a topological space $X$ *converges weakly to a Borel measure* $\mu$ *on* $X$, if for every continuous and bounded function $f\colon X \to \mathbb{R}$ we have

$$\lim_{n\to\infty} \int f \, \mathrm{d}\mu_n = \int f \, \mathrm{d}\mu;$$

see, e.g., [Billingsley 1968, Section 1.1].

**Lemma 2.3.** *Let $(\mathfrak{D}_n)_{n\in\mathbb{N}}$ be a sequence of effective divisors on $\mathbb{C}_p$ such that for every $n$ we have* $\deg(\mathfrak{D}_n) \ge 1$. *Then, the following are equivalent*:

(i) $\bar{\delta}_{\iota(\mathfrak{D}_n)} \to \delta_{x_{\mathrm{can}}}$ *weakly as* $n \to \infty$.

(ii) *For every $R > 1$ and every $a$ in $\mathcal{O}_p$, we have for $\boldsymbol{D} = \boldsymbol{D}(a, R^{-1})$ and $\boldsymbol{D} = \boldsymbol{D}^\infty(a, R)$,*

$$\lim_{n\to\infty} \frac{\deg(\mathfrak{D}_n|_{\boldsymbol{D}})}{\deg(\mathfrak{D}_n)} = \lim_{n\to\infty} \bar{\delta}_{\mathfrak{D}_n}(\boldsymbol{D}) = 0.$$

For the reader's convenience we provide a self-contained proof of this lemma, which applies to the Berkovich affine line over an arbitrary complete and algebraically closed field. Using that $\mathbb{A}^1_{\mathrm{Berk}}$ is

metrizable, the lemma can also be obtained as a direct consequence of the following observations: (i) is equivalent to the assertion that for every neighborhood $\mathcal{U}$ of $x_{\mathrm{can}}$ in $\mathbb{A}^1_{\mathrm{Berk}}$ we have

$$\lim_{n\to\infty} \bar{\delta}_{\mathfrak{D}_n}(\mathcal{U}) = 1.$$

This last statement is equivalent to the contrapositive of (ii).

*Proof of Lemma 2.3.* Assume that (i) holds and let $R > 1$ and $a$ in $\mathcal{O}_p$ be given. Note that the first equality in (ii) is a direct consequence of the definitions. To prove the second equality, take a continuous function $\phi\colon \mathbb{R}_0^+ \to [0, 1]$ satisfying $\phi(1) = 0$ and $\phi(t) = 1$ for $0 \le t \le R^{-1}$ and for $t \ge R$. Let $\alpha\colon \mathbb{A}^1_{\mathrm{Berk}} \to \mathbb{R}$ be the continuous function given by $\alpha(x) = x(X - a)$ and put $F := \phi \circ \alpha$. By construction we have

$$F(x_{\mathrm{can}}) = \phi(1) = 0 \quad \text{and} \quad F(x) = 1 \text{ for all } x \in \mathcal{D}(a, R^{-1}) \cup \mathcal{D}^\infty(a, R).$$

Using that for $z \in \mathbb{C}_p$ we have

$$z \in \boldsymbol{D}(a, R^{-1}) \Leftrightarrow \iota(z) \in \mathcal{D}(a, R^{-1}) \quad \text{and} \quad z \in \boldsymbol{D}^\infty(a, R) \Leftrightarrow \iota(z) \in \mathcal{D}^\infty(a, R), \tag{2-19}$$

we get

$$0 \le \bar{\delta}_{\mathfrak{D}_n}(\boldsymbol{D}(a, R^{-1}) \cup \boldsymbol{D}^\infty(a, R)) = \bar{\delta}_{\iota(\mathfrak{D}_n)}(\mathcal{D}(a, R^{-1}) \cup \mathcal{D}^\infty(a, R)) \le \int F \, \mathrm{d}\bar{\delta}_{\iota(\mathfrak{D}_n)}.$$

Since $F$ is continuous and bounded, our hypothesis (i) implies that

$$\bar{\delta}_{\mathfrak{D}_n}(\boldsymbol{D}(a, R^{-1})) \to 0 \quad \text{and} \quad \bar{\delta}_{\mathfrak{D}_n}(\boldsymbol{D}^\infty(a, R)) \to 0 \text{ as } n \to \infty.$$

This completes the proof of the implication (i) $\Rightarrow$ (ii).

Now, assume that (ii) holds, let $F\colon \mathbb{A}^1_{\mathrm{Berk}} \to \mathbb{R}$ be a continuous and bounded function and let $\varepsilon > 0$ be given. Since the sets (2-18) form a basis of neighborhoods of $x_{\mathrm{can}}$, there are $R > 1$ and a finite subset $A$ of $\mathcal{O}_p$ such that

$$|F(x) - F(x_{\mathrm{can}})| < \varepsilon \quad \text{for all } x \in \mathcal{A}(A; R). \tag{2-20}$$

Let $R'$ in $]1, R[$ be fixed. From the definition of $\mathcal{A} := \mathcal{A}(A; R)$, we have

$$\mathcal{A}' := \mathbb{A}^1_{\mathrm{Berk}} \setminus \mathcal{A} \subseteq \mathcal{D}^\infty(0, R') \cup \bigcup_{a\in A} \mathcal{D}(a, (R')^{-1}).$$

Using (2-19) and (ii) with $R$ replaced by $R'$ and with $a$ in $A \cup \{0\}$, we obtain

$$\begin{aligned}
\deg(\iota(\mathfrak{D}_n)|_{\mathcal{A}'}) &\le \deg(\iota(\mathfrak{D}_n)|_{\mathcal{D}^\infty(0,R')}) + \sum_{a\in A} \deg(\iota(\mathfrak{D}_n)|_{\mathcal{D}(a,(R')^{-1})}) \\
&= \deg(\mathfrak{D}_n|_{\boldsymbol{D}^\infty(0,R')}) + \sum_{a\in A} \deg(\mathfrak{D}_n|_{\boldsymbol{D}(a,(R')^{-1})}) \\
&= o(\deg(\iota(\mathfrak{D}_n))).
\end{aligned}$$

Together with our choice of $\mathcal{A}(A; R)$, this implies

$$\begin{aligned}\left|\int F \, \mathrm{d}\bar{\delta}_{\iota(\mathfrak{D}_n)} - F(x_{\mathrm{can}})\right| &\\ \leq \left|\frac{F(\iota(\mathfrak{D}_n)|_{\mathcal{A}}) - F(x_{\mathrm{can}})\deg(\iota(\mathfrak{D}_n)|_{\mathcal{A}})}{\deg(\mathfrak{D}_n)}\right| &+ \left|\frac{F(\iota(\mathfrak{D}_n)|_{\mathcal{A}'}) - F(x_{\mathrm{can}})\deg(\iota(\mathfrak{D}_n)|_{\mathcal{A}'})}{\deg(\mathfrak{D}_n)}\right| \\ \leq \varepsilon + 2(\sup_{x\in\mathbb{A}^1_{\mathrm{Berk}}} |F(x)|)&\frac{\deg(\iota(\mathfrak{D}_n)|_{\mathcal{A}'})}{\deg(\iota(\mathfrak{D}_n)},\end{aligned}$$

and therefore

$$\limsup_{n\to\infty}\left|\int F \, \mathrm{d}\bar{\delta}_{\iota(\mathfrak{D}_n)} - F(x_{\mathrm{can}})\right| \leq \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, this completes the proof of the implication (ii) $\Rightarrow$ (i) and of the lemma. □

## 3. CM points in the ordinary reduction locus

The purpose of this section is to give a strengthened version of Theorem A in the case where all the discriminants in the sequence $(D_n)_{n=1}^{\infty}$ are $p$-ordinary (Theorem 3.5(ii) in Section 3B). An important tool is "the canonical branch $\boldsymbol{t}$" of $T_p$ on $Y_{\mathrm{ord}}(\mathbb{C}_p)$, which is defined using the canonical subgroup in Section 3A. We use it to give, for every integer $m \geq 1$, a simple formula of $T_{p^m}$ (Proposition 3.4 in Section 3A). Moreover, we show that $p$-ordinary CM points correspond precisely to the preperiodic points of $\boldsymbol{t}$ on $Y_{\mathrm{ord}}(\mathbb{C}_p)$ (Theorem 3.5(i)). Once these are established, Theorem 3.5(ii) follows from dynamical properties of $\boldsymbol{t}$ on $Y_{\mathrm{ord}}(\mathbb{C}_p)$ (Lemma 3.7). In Appendix B we extend and further study the canonical branch $\boldsymbol{t}$ of $T_p$.

We use properties of reduction morphisms that are stated in most of the classical literature only for elliptic curves over discrete valued fields. To extend the application of these results to elliptic curves over $\mathbb{C}_p$ we use the continuity of the Hecke correspondences (Lemma 2.1 in Section 2B). To this purpose, we introduce the following notation: $\mathbb{Q}_p^{\mathrm{unr}}$ is the maximal unramified extension of $\mathbb{Q}_p$ inside $\overline{\mathbb{Q}}_p$, and $\mathbb{C}_p^{\mathrm{unr}}$ is its completion. Then, $\mathbb{C}_p^{\mathrm{unr}}$ is an infinite degree extension of $\mathbb{Q}_p$ with the same valuation group and with residue field $\bar{\mathbb{F}}_p$. The algebraic closure $\overline{\mathbb{C}_p^{\mathrm{unr}}}$ of $\mathbb{C}_p^{\mathrm{unr}}$ inside $\mathbb{C}_p$ is dense in $\mathbb{C}_p$. Since $\overline{\mathbb{C}_p^{\mathrm{unr}}}$ can be written as the union of finite extensions of $\mathbb{C}_p^{\mathrm{unr}}$, it follows that every elliptic curve in $Y(\overline{\mathbb{C}_p^{\mathrm{unr}}})$ can be defined over a complete discrete valued field with residue field $\bar{\mathbb{F}}_p$. The same holds for finite subgroups and isogenies between elliptic curves over $\overline{\mathbb{C}_p^{\mathrm{unr}}}$.

In what follows, we use $Y_{\mathrm{ord}}(\overline{\mathbb{C}_p^{\mathrm{unr}}}) := Y_{\mathrm{ord}}(\mathbb{C}_p) \cap Y(\overline{\mathbb{C}_p^{\mathrm{unr}}})$.

**3A. *The canonical branch of $T_p$ on $Y_{\mathrm{ord}}(\mathbb{C}_p)$.*** In this section we define a branch of the Hecke correspondence $T_p$ on $Y_{\mathrm{ord}}(\mathbb{C}_p)$ that we use to give a simple description, for every integer $m \geq 1$, of $T_{p^m}$ that is crucial in what follows (Proposition 3.4). See also Appendix B. We start recalling the following result describing the endomorphism ring of the reduction of a CM point in the ordinary locus.

**Proposition 3.1** [Lang 1973, Chapter 13, Section 4, Theorem 12]. *Let $d<0$ be a fundamental discriminant and let $f \geq 1$ and $m \geq 0$ be integers such that $f$ is not divisible by $p$. Then, for an elliptic curve $E$ defined over a discrete valued subfield of $\mathbb{C}_p$ having ordinary reduction,* $\mathrm{End}(E) \simeq \mathcal{O}_{d,p^m f}$ *implies that the reduction $\tilde{E}$ of $E$ satisfies* $\mathrm{End}(\tilde{E}) \simeq \mathcal{O}_{d,f}$. *In particular, if* $\mathrm{End}(E)$ *is an order in a quadratic imaginary extension of $\mathbb{Q}$ whose conductor is not divisible by $p$, then the reduction map* $\mathrm{End}(E) \to \mathrm{End}(\tilde{E})$ *is an isomorphism.*

To define the canonical branch of $T_p$ on $Y_{\mathrm{ord}}(\mathbb{C}_p)$, we use the *canonical subgroup* of an elliptic curve $E$ in $Y_{\mathrm{ord}}(\overline{\mathbb{C}_p^{\mathrm{unr}}})$, which is defined as the unique subgroup of order $p$ of $E$ in the kernel of the reduction morphism $E \to \tilde{E}$. Equivalently, $H(E)$ is the kernel of the reduction morphism $E[p] \to \tilde{E}[p]$. For an elliptic curve $e \in Y(\bar{\mathbb{F}}_p)$ denote by $\mathrm{Frob}\colon e \to e^{(p)}$ the Frobenius morphism, which is the isogeny given in affine coordinates by $(x, y) \mapsto (x^p, y^p)$.

**Theorem 3.2.** (i) *For $E$ in $Y_{\mathrm{ord}}(\overline{\mathbb{C}_p^{\mathrm{unr}}})$ the reduction of $E/H(E)$ equals $\tilde{E}^{(p)}$ and every isogeny $\varphi\colon E \to E/H(E)$ whose kernel is equal to $H(E)$ reduces to the Frobenius morphism*

$$\mathrm{Frob}\colon \tilde{E} \to \tilde{E}^{(p)}.$$

*Moreover, the kernel of the isogeny dual to $\varphi$ is different from the canonical subgroup of $E/H(E)$.*

(ii) *For each ordinary elliptic curve $e \in Y(\bar{\mathbb{F}}_p)$ there exists a unique elliptic curve $e^{\uparrow} \in Y(\overline{\mathbb{C}_p^{\mathrm{unr}}})$ reducing to $e$ for which the reduction map induces a ring isomorphism*

$$\mathrm{End}(e^{\uparrow}) \simeq \mathrm{End}(e).$$

(iii) *Given two ordinary elliptic curves $e_1, e_2 \in Y(\bar{\mathbb{F}}_p)$, the reduction map induces a group isomorphism*

$$\mathrm{Hom}(e_1^{\uparrow}, e_2^{\uparrow}) \simeq \mathrm{Hom}(e_1, e_2).$$

*In particular, the Frobenius morphism* $\mathrm{Frob}\colon e \to e^{(p)}$ *lifts to an isogeny $e^{\uparrow} \to (e^{(p)})^{\uparrow}$ with kernel $H(e^{\uparrow})$, and $e^{\uparrow}/H(e^{\uparrow}) = (e^{(p)})^{\uparrow}$.*

*Proof.* Item (i) follows from the definition of canonical subgroup and properties of reduction morphisms; see, e.g., [Diamond and Shurman 2005, Proof of Lemma 8.7.1]. Item (ii) is usually known as "Deuring's lifting theorem", see for example [Deuring 1941] or [Lang 1973, Chapter 13, Section 5, Theorem 14]. Item (iii) is another known consequence of Deuring's work. To prove surjectivity, first note that every isogeny in $\mathrm{Hom}(e_1, e_2)$ can be written as a composition of Frobenius morphisms, of duals of Frobenius morphisms, and of an isogeny whose degree is not divisible by $p$. In view of items (i) and (ii), and of Proposition 3.1, we can restrict to the case of an isogeny of degree $n$ not divisible by $p$. This case is a direct consequence of item (ii), and the fact that the reduction morphism $E \to \tilde{E}$ induces a bijective map $E[n] \to \tilde{E}[n]$, see for example [Silverman 2009, Chapter VII, Proposition 3.1(b)]. □

The following result is due to Tate in the case $p = 2$ and to Deligne in the general case. To state it, define

$$\begin{aligned} \boldsymbol{t}\colon Y_{\mathrm{ord}}(\overline{\mathbb{C}_p^{\mathrm{unr}}}) &\to Y_{\mathrm{ord}}(\overline{\mathbb{C}_p^{\mathrm{unr}}}) \\ E &\mapsto \boldsymbol{t}(E) := E/H(E), \end{aligned} \tag{3-1}$$

and for $e$ in $Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)$ put

$$\delta_e := \begin{cases} 1 & \text{if } p \geq 5,\ j(e) \neq 0, 1728; \\ 3 & \text{if } p \geq 5,\ j(e) = 0; \\ 2 & \text{if } p \geq 5,\ j(e) = 1728; \\ 6 & \text{if } p = 3,\ j(e) = 0 = 1728; \\ 12 & \text{if } p = 2,\ j(e) = 0 = 1728. \end{cases} \tag{3-2}$$

Note that in all the cases $\delta_e = (\#\operatorname{Aut}(e))/2$; see, e.g., [Silverman 1994, Chapter III, Theorem 10.1].

**Theorem 3.3.** *For each $e$ in $Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)$ choose $\beta_e$ in $\boldsymbol{D}(j(e)) \cap \mathbb{Q}_p^{\mathrm{unr}}$, so that $\pi(\beta_e) = j(e)$, and put $\delta'_e := \delta_e$ if $\beta_e = 0$ and $p \neq 3$ or if $\beta_e = 1728$ and $p \neq 2$, and $\delta'_e := 1$ otherwise. Then, the map $\boldsymbol{t}$ admits an expansion of the form*

$$\boldsymbol{t}(z) = z^p + pk(z) + \sum_{e \in Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)} \sum_{n=1}^{\infty} \frac{A_n^{(e)}}{(z - \beta_e)^n}, \tag{3-3}$$

*where $k(z)$ is a polynomial of degree $p-1$ in $z$ with coefficients in $\mathbb{Z}$, and for each $n \geq 1$ the coefficient $A_n^{(e)}$ belongs to $\mathbb{Q}_p(\{\beta_e : e \in Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)\})$ and*

$$\operatorname{ord}_p(A_n^{(e)}) \geq \delta'_e \left( \frac{1}{p+1} + n \frac{p}{p+1} \right). \tag{3-4}$$

*In particular, $\boldsymbol{t}(z)$ extends to a rigid analytic function $Y_{\mathrm{ord}}(\mathbb{C}_p) \to Y_{\mathrm{ord}}(\mathbb{C}_p)$ of degree $p$ that we also denote by $\boldsymbol{t}$.*

For $p \geq 5$, this result is proved in [Dwork 1969, Chapter 7]. In the case $\delta'_e > 1$, (3-4) can be obtained from the method of proof described in [loc. cit.], or from the estimate in [loc. cit., page 80] combined with the fact that $\operatorname{ord}_p(A_n^{(e)})$ is an integer and that $\beta_e = 0$ implies $p \equiv 2 \bmod 3$. For $p = 2$ and 3, this result is stated in [loc. cit., page 89] with a weaker version of (3-4). We provide the details of the proof when $p = 2$ and 3; see Proposition B.2 in Appendix B.

The theorem above implies that $\boldsymbol{t}$ extends to a rigid analytic map from $Y_{\mathrm{ord}}(\mathbb{C}_p)$ to itself. We denote this extension also by $\boldsymbol{t}$ and call it the *canonical branch of $T_p$ on $Y_{\mathrm{ord}}(\mathbb{C}_p)$*.

For $z \in Y_{\mathrm{ord}}(\mathbb{C}_p)$, let $\boldsymbol{t}^*(z)$ be the divisor on $Y_{\mathrm{ord}}(\mathbb{C}_p)$ given by

$$\boldsymbol{t}^*(z) := \sum_{\substack{w \in Y_{\mathrm{ord}}(\mathbb{C}_p) \\ \boldsymbol{t}(w) = z}} \deg_{\boldsymbol{t}}(w)[w],$$

where $\deg_{\boldsymbol{t}}(w)$ is the local degree of $\boldsymbol{t}$ at $w$. Note that by Theorem 3.3 the rigid analytic map $\boldsymbol{t}\colon Y_{\mathrm{ord}}(\mathbb{C}_p)\to Y_{\mathrm{ord}}(\mathbb{C}_p)$ is of degree $p$, so for $z$ in $Y_{\mathrm{ord}}(\mathbb{C}_p)$ we have

$$\deg(\boldsymbol{t}^*(z))=p \quad \text{and} \quad \boldsymbol{t}_*(\boldsymbol{t}^*(z))=p[z].$$

As usual, for an integer $i\geq 1$ we denote by $\boldsymbol{t}^i$ the $i$-th fold composition of $\boldsymbol{t}$ with itself. We also use $\boldsymbol{t}^0$ to denote the identity on $Y_{\mathrm{ord}}(\mathbb{C}_p)$.

**Proposition 3.4.** *For every $E$ in $Y_{\mathrm{ord}}(\mathbb{C}_p)$ and every integer $m\geq 1$, we have*

$$T_{p^m}(E)=\sum_{i=0}^{m}(\boldsymbol{t}^*)^{m-i}([\boldsymbol{t}^i(E)]). \tag{3-5}$$

When $m=1$, the relation (3-5) reads

$$T_p(E)=\boldsymbol{t}^*(E)+[\boldsymbol{t}(E)]. \tag{3-6}$$

See Theorem B.1 in Appendix B for an extension.

*Proof.* The relation (3-5) for $m\geq 2$ follows from (3-6) by induction using the recursive formula (2-5). To prove (3-6), first note that for $E$ in $Y_{\mathrm{ord}}(\mathbb{C}_p)$ satisfying $\deg_{\boldsymbol{t}}(E)\geq 2$ we have $\boldsymbol{t}'(E)=0$. Therefore there are at most a finite number of such $E$ in the affinoid $Y_{\mathrm{ord}}(\mathbb{C}_p)$; see for example [Fresnel and van der Put 2004, Proposition 3.3.6]. It follows that for every $E$ in $Y_{\mathrm{ord}}(\mathbb{C}_p)$ outside a finite set of exceptions, we have $\#\operatorname{supp}(\boldsymbol{t}^*(E))=p$. Thus, the set D of all those $E$ in $Y_{\mathrm{ord}}(\overline{\mathbb{C}_p^{\mathrm{unr}}})$ with this property is dense in $Y_{\mathrm{ord}}(\mathbb{C}_p)$. To prove (3-6) for $E$ in D, use the definition of $T_p(E)$ and $\boldsymbol{t}(E)$, and Theorem 3.2(i), to obtain

$$T_p(E)=[\boldsymbol{t}(E)]+\sum_{\substack{C\leq E,\#C=p\\ C\neq H(E)}}[E/C]=[\boldsymbol{t}(E)]+\boldsymbol{t}^*(E).$$

To prove (3-6) for an arbitrary $E$ in $Y_{\mathrm{ord}}(\mathbb{C}_p)$, first note that by Lemma 2.1 for every open and closed subset $A$ of $Y_{\mathrm{ord}}(\mathbb{C}_p)$ the function

$$E\mapsto \mathbf{1}_A(T_n(E)-\boldsymbol{t}^*(E)-[\boldsymbol{t}(E)])=\deg((T_n(E)-\boldsymbol{t}^*(E)-[\boldsymbol{t}(E)])|_A)$$

is continuous. Since it is equal to 0 on the dense subset D of $Y_{\mathrm{ord}}(\mathbb{C}_p)$, we conclude that it is constant equal to 0. Since this holds for every open and closed subset $A$ of $Y_{\mathrm{ord}}(\mathbb{C}_p)$, this proves (3-6) and completes the proof of the lemma. □

**3B.** ***CM points as preperiodic points.*** The purpose of this section is to prove the following result. In the case where all the discriminants in the sequence $(D_n)_{n=1}^{\infty}$ are $p$-ordinary, Theorem A is a direct consequence of item (ii) of this result together with (2-8) and Lemma 2.3.

Given a set $X$ and a map $T\colon X\to X$, a point $x$ in $X$ is *periodic* if for some integer $r\geq 1$ we have $T^r(x)=x$. Then the integer $r$ is a *period of* $x$ and the smallest such integer is the *minimal period of* $x$. Moreover, a point $y$ is *preperiodic* if it is not periodic and if for some integer $m\geq 1$ the point $T^m(y)$ is periodic. We call the least such integer $m$ the *preperiod of* $y$.

**Theorem 3.5.** *Let $\zeta$ in $\bar{\mathbb{F}}_p$ be the $j$-invariant of an ordinary elliptic curve and denote by $r$ the minimal period of $\zeta$ under the Frobenius map $z \mapsto z^p$. Then there is a unique periodic point $E_0$ of $\boldsymbol{t}$ in $\boldsymbol{D}(\zeta)$. The minimal period of $E_0$ is $r$. Moreover, $E_0$ is a CM point and, if we denote by $D_0$ the discriminant of the endomorphism ring of $E_0$, then the conductor of $D_0$ is not divisible by $p$ and the following properties hold*:

(i) *Given a discriminant $D$, the set* $\operatorname{supp}(\Lambda_D|_{\boldsymbol{D}(\zeta)})$ *is nonempty if and only if for some integer $m \geq 0$ we have $D = D_0 p^{2m}$. Moreover,*

$$\operatorname{supp}(\Lambda_{D_0}|_{\boldsymbol{D}(\zeta)}) = \{E_0\}$$

*and for each integer $m \geq 1$ the set* $\operatorname{supp}(\Lambda_{D_0 p^{2m}}|_{\boldsymbol{D}(\zeta)})$ *is equal to the set of all the preperiodic points of $\boldsymbol{t}$ in $\boldsymbol{D}(\zeta)$ of preperiod $m$, and is contained in $\boldsymbol{t}^{-m}(\boldsymbol{t}^m(E_0))$. In particular, CM points in $Y_{\mathrm{ord}}(\mathbb{C}_p)$ correspond precisely to the periodic and preperiodic points of $\boldsymbol{t}$ in $Y_{\mathrm{ord}}(\mathbb{C}_p)$.*

(ii) *For every disc $\boldsymbol{B}$ of radius strictly less than 1 contained in $\boldsymbol{D}(\zeta)$ there is a constant $C > 0$ such that for every discriminant $D < 0$, we have*

$$\deg(\Lambda_D|_{\boldsymbol{B}}) \leq C.$$

**Remark 3.6.** The natural directed graph associated to the dynamics of $\boldsymbol{t}$ on the set of ordinary CM points is a "$(p+1)$-volcano" in the sense of [Goren and Kassaei 2017, Section 2.1]. This follows from Theorem 3.5(i) and the fact that $\boldsymbol{t}$ is of degree $p$ on $Y_{\mathrm{ord}}(\mathbb{C}_p)$ by Theorem 3.3. Note in particular that the "rim" is the directed subgraph associated to the dynamics of $\boldsymbol{t}$ on the set of its periodic points in $Y_{\mathrm{ord}}(\mathbb{C}_p)$. Moreover, on the set of preperiodic points of $\boldsymbol{t}$ in $Y_{\mathrm{ord}}(\mathbb{C}_p)$, the preperiod corresponds to the function "$b$" of [Goren and Kassaei 2017].

To prove Theorem 3.5, we describe the dynamics of $\boldsymbol{t}$ on $Y_{\mathrm{ord}}(\mathbb{C}_p)$ in Lemma 3.7 below. This description is mostly based on the fact that

$$\boldsymbol{t}(z) \equiv z^p \bmod p\mathcal{O}_p, \tag{3-7}$$

see Theorem 3.3. We deduce from general considerations that each residue disc $\boldsymbol{D} \subseteq Y_{\mathrm{ord}}(\mathbb{C}_p)$ contains a unique periodic point $z_0$ of $\boldsymbol{t}$, that this point satisfies $|\boldsymbol{t}'(z_0)| < 1$, and that every point in $\boldsymbol{D}$ is asymptotic to $z_0$.[‡] The fact that no periodic point of $\boldsymbol{t}$ in $Y_{\mathrm{ord}}(\mathbb{C}_p)$ is a ramification point is used in a crucial way in the proof of the estimate (5-5) of Proposition 5.3 in Section 5B.

**Lemma 3.7** (dynamics of $\boldsymbol{t}$ on $Y_{\mathrm{ord}}(\mathbb{C}_p)$). *Let $e$ be an ordinary elliptic curve defined over $\bar{\mathbb{F}}_p$ and let $r \geq 1$ be the minimal period of $j(e)$ under the Frobenius map. Then, $e^{\uparrow}$ is the unique elliptic curve in $\boldsymbol{D}(j(e))$ that is periodic for $\boldsymbol{t}$. The minimal period of $e^{\uparrow}$ for $\boldsymbol{t}$ is $r$ and $e^{\uparrow}$ is also characterized as the unique elliptic curve in $\boldsymbol{D}(j(e)) \cap \overline{\mathbb{C}_p}^{\mathrm{unr}}$ whose endomorphism ring is an order in an quadratic imaginary extension of $\mathbb{Q}$ of conductor not divisible by $p$. Moreover, if for every integer $i \geq 0$ we put $z_i := \boldsymbol{t}^i(e^{\uparrow})$, then the following properties hold*:

[‡]This is somewhat similar to the case of a rational map having good reduction equal to the Frobenius map, see for example [Rivera-Letelier 2003, Sections 3.1 and 4.5].

(i) *For each integer $i \geq 0$ we have $0 < |\boldsymbol{t}'(z_i)|_p < 1$.*

(ii) *There is $\rho$ in $]0, 1[$ such that for every integer $i \geq 0$ and all $z$ and $z'$ in $\boldsymbol{D}(z_i, \rho)$, we have*

$$\deg_{\boldsymbol{t}}(z) = 1 \quad \textit{and} \quad |\boldsymbol{t}(z) - \boldsymbol{t}(z')|_p = |\boldsymbol{t}'(z_i)|_p \cdot |z - z'|_p.$$

*In particular, $\boldsymbol{t}$ is injective on $\boldsymbol{D}(z_i, \rho)$.*

(iii) *For every $c \in ]0, 1[$ there exists $\kappa_c$ in $]0, 1[$ such that for every integer $i \geq 0$, every $z$ in $\boldsymbol{D}(z_i, 1)$ satisfying $|z - z_i|_p \leq c$ and every integer $m \geq 1$, we have*

$$|\boldsymbol{t}^m(z) - z_{i+m}|_p \leq \kappa_c^m |z - z_i|_p.$$

(iv) *For all $i \geq 0$ and $z$ in $\boldsymbol{D}(z_i, 1)$, the sequence*

$$(|\boldsymbol{t}^m(z) - z_{i+m})|_p)_{m=0}^{\infty}$$

*is nonincreasing and converges to 0.*

*Proof.* We start by proving (i). Suppose for a contradiction that $z_i$ is a ramification point of $\boldsymbol{t}$. Without loss of generality, assume that $i = 0$ and put $E := e^{\uparrow}$ and $E^p := (e^{(p)})^{\uparrow}$. By Proposition 3.4 with $m = 1$ there are distinct subgroups $C$ and $C'$ of $E^p$ of order $p$ such that

$$E^p/C = E^p/C' = E, C \neq H(E^p) \quad \text{and} \quad C' \neq H(E^p).$$

Let $\psi$ (resp. $\psi'$) be an isogeny $E^p \to E$ with kernel $C$ (resp. $C'$) and denote by $\hat{\psi}$ (resp. $\hat{\psi}'$) its dual isogeny. Then the kernel of $\hat{\psi}$ and of $\hat{\psi}'$ are both equal to $H(E)$. It follows that there is $\sigma$ in $\mathrm{Aut}(E^p)$ such that $\sigma \circ \hat{\psi} = \hat{\psi}'$; see, e.g., [Silverman 2009, Chapter III, Corollary 4.11]. Since $\sigma \neq \pm 1$, we have $j(E^p) \in \{0, 1728\}$ and therefore $r = 1$, $\boldsymbol{t}(z_0) = z_0$ and $E^p = E$. In particular, $C$ and $C'$ are subgroups of $E$ and $\psi, \psi' \in \mathrm{End}(E)$. The kernel of each of the reduced isogenies $\tilde{\psi}$ and $\tilde{\psi}'$ is equal to $e[p](\bar{\mathbb{F}}_p)$, so there is $\tilde{\alpha}$ in $\mathrm{Aut}(e)$ such that $\tilde{\alpha} \circ \tilde{\psi} = \tilde{\psi}'$. Since the reduction map $\mathrm{End}(E) \to \mathrm{End}(e)$ is an isomorphism by Theorem 3.2(ii), we can find an automorphism $\alpha \in \mathrm{Aut}(E)$ satisfying $\alpha \circ \psi = \psi'$. This implies that the kernel $C$ of $\psi$ is equal to the kernel $C'$ of $\psi'$, and we obtain a contradiction. This completes the proof that $z_i$ is not a ramification point of $\boldsymbol{t}$ and therefore that $\boldsymbol{t}'(z_i) \neq 0$.

To prove that $|\boldsymbol{t}'(z_i)| < 1$ note that by Theorem 3.3, we can write

$$\boldsymbol{t}(w + z_i) - z_{i+1} = \boldsymbol{t}(w + z_i) - \boldsymbol{t}(z_i) = \sum_{n=1}^{\infty} B_n^{(i)} w^n, \tag{3-8}$$

where the coefficients $B_n^{(i)}$ belong to $\mathcal{O}_p$ and satisfy $|B_n^{(i)}|_p \leq \frac{1}{p}$ for $n \neq p$. Since $\boldsymbol{t}'(z_i) = B_1^{(i)}$, this completes the proof of (i).

To prove the assertions at the beginning of the lemma, for each integer $i \geq 0$ denote by $e^{(p^i)}$ the image of $e$ by the $i$-th iterate of the Frobenius morphism. Then by Theorem 3.2(iii) we have

$$z_i = \boldsymbol{t}^i(e^{\uparrow}) = (e^{(p^i)})^{\uparrow} \in \pi^{-1}(j(e)^{p^i}).$$

It follows that $z_0$ is periodic of minimal period $r$ for $\boldsymbol{t}$. To prove uniqueness, note that by (3-8) for every integer $i \geq 0$ and distinct $z$ and $z'$ in $D(z_i, 1)$ we have

$$|\boldsymbol{t}(z) - \boldsymbol{t}(z')|_p < |z - z'|_p. \tag{3-9}$$

Thus, there can be at most one periodic point of $\boldsymbol{t}$ in $\boldsymbol{D}(z_0, 1)$. Finally, combining Theorem 3.2(ii) and Proposition 3.1 we obtain that $e^{\uparrow}$ is the unique elliptic curve reducing to $e$ and whose endomorphism ring is an order of conductor not divisible by $p$. This completes the proof of the assertions at the beginning of the proposition, so it only remains to prove (ii), (iii) and (iv).

To prove (ii), note that by (i) there is $\rho$ in $]0, 1[$ so that for every $i$ in $\{0, \ldots, r-1\}$, we have

$$\max\{|B_n^{(i)}|_p \rho^{n-1} : n \geq 2\} \leq |B_1^{(i)}|_p.$$

Then by the ultrametric inequality for every integer $i \geq 0$ and $z \in \boldsymbol{D}(z_i, \rho)$ we have $|\boldsymbol{t}'(z)|_p = |B_1^{(i)}|_p$, which is different from 0 by (i). In particular, $\deg_{\boldsymbol{t}}(z_i) = 1$. Moreover, for $z'$ in $\boldsymbol{D}(z_i, \rho)$ we have by the ultrametric inequality

$$|\boldsymbol{t}(z) - \boldsymbol{t}(z')|_p = |B_1^{(i)}|_p |z - z'|_p.$$

This completes the proof of (ii).

Item (iii) is a direct consequence of (3-8) with

$$\kappa_c := \max\{|B_n^{(i)}| c^{n-1} : n \geq 1, i \in \{0, \ldots, r-1\}\},$$

noting that for every integer $n \geq 1$ and all integers $i, i' \geq 0$ such that $i - i'$ is divisible by $r$, we have $B_n^{(i')} = B_n^{(i)}$.

To prove item (iv), note that the fact that the sequence is nonincreasing follows from (3-9) and the fact that it converges to 0 form (iii) with $c = |z - z_i|$. This completes the proof the lemma. □

*Proof of Theorem 3.5.* The first assertions are given by Lemma 3.7.

To prove (i), note that Proposition 3.1 implies that if a discriminant $D < 0$ is such that $\operatorname{supp}(\Lambda_D|_{\boldsymbol{D}(\zeta)})$ is nonempty, then there is an integer $m \geq 0$ such that $D = D_0 p^{2m}$. On the other hand, Lemma 3.7 implies $\operatorname{supp}(\Lambda_{D_0}|_{\boldsymbol{D}(\zeta)}) = \{E_0\}$. Fix an integer $m \geq 1$ and note that by Lemma 3.7 for every integer $j \geq 1$ the point $E_j := \boldsymbol{t}^j(E_0)$ is the unique periodic point of $\boldsymbol{t}$ in $\boldsymbol{D}(\zeta^{p^j})$. So, if $E$ is a preperiodic point of $\boldsymbol{t}$ in $\boldsymbol{D}(\zeta)$ of preperiod $m$, then $\boldsymbol{t}^m(E) = E_m$. This implies that the set of all preperiodic points of $\boldsymbol{t}$ in $\boldsymbol{D}(\zeta)$ of preperiod $m$ is contained in $\boldsymbol{t}^{-m}(E_m)$ and is equal to

$$\boldsymbol{t}^{-m}(E_m) \setminus \boldsymbol{t}^{-(m-1)}(E_{m-1}) = \boldsymbol{t}^{-(m-1)}(\boldsymbol{t}^{-1}(E_m) \setminus \{E_{m-1}\}).$$

Since the degree of $\boldsymbol{t}$ is $p$ and by Lemma 3.7(i) we have $\boldsymbol{t}'(E_{m-1}) \neq 0$, the set $\boldsymbol{t}^{-1}(E_m) \setminus \{E_{m-1}\}$ is nonempty and equal to $\operatorname{supp}(\boldsymbol{t}^*([E_m]) - [E_{m-1}])$. We thus conclude that the set of preperiodic points of $\boldsymbol{t}$ in $\boldsymbol{D}(\zeta)$ of preperiod $m$ is equal to $\boldsymbol{t}^{-(m-1)}(\operatorname{supp}(\boldsymbol{t}^*([E_m]) - [E_{m-1}]))$ and it is nonempty. Thus, to complete the proof of (i) it is sufficient to show that the set of preperiodic points of $\boldsymbol{t}$ in $\boldsymbol{D}(\zeta)$ of preperiod

$m$ is equal to $\mathrm{supp}(\Lambda_{D_0 p^{2m}}|_{\boldsymbol{D}(\zeta)})$. Note that by (2-11) and Proposition 3.4 we have

$$\mathrm{supp}(\boldsymbol{t}_*(\Lambda_{D_0})) \subseteq \mathrm{supp}(T_p(\Lambda_{D_0})) = \mathrm{supp}(\Lambda_{D_0}) \cup \mathrm{supp}(\Lambda_{D_0 p^2}).$$

By Lemma 3.7 the set $\mathrm{supp}(\Lambda_{D_0})$, hence $\mathrm{supp}(\boldsymbol{t}_*(\Lambda_{D_0}))$, is formed by periodic points of $\boldsymbol{t}$ while points in $\mathrm{supp}(\Lambda_{D_0 p^2})$ are not periodic. This implies

$$\boldsymbol{t}_*(\Lambda_{D_0}) = \Lambda_{D_0}. \tag{3-10}$$

Let $d$ and $f_0$ be the fundamental discriminant and conductor of $D_0$, respectively. Since $p$ splits in $\mathbb{Q}(\sqrt{d})$ we deduce that for every integer $k \geq 0$ we have $R_d(p^k) = k+1$. By (2-9), Proposition 3.4 and (3-10) we get

$$\mathrm{supp}((\boldsymbol{t}^*)^m(\Lambda_{D_0})) = \bigcup_{k=0}^{m} \mathrm{supp}(\Lambda_{D_0 p^{2k}}).$$

This implies the equality

$$\mathrm{supp}((\boldsymbol{t}^*)^m(\Lambda_{D_0})) \setminus \mathrm{supp}((\boldsymbol{t}^*)^{m-1}(\Lambda_{D_0})) = \mathrm{supp}(\Lambda_{D_0 p^{2m}}). \tag{3-11}$$

By Lemma 3.7 and (3-10) the set $\mathrm{supp}(\Lambda_{D_0}) \cap (\boldsymbol{D}(\zeta) \cup \boldsymbol{D}(\zeta^p) \cup \cdots \cup \boldsymbol{D}(\zeta^{p^{r-1}}))$ equals the set of periodic points of $\boldsymbol{t}$ in $\boldsymbol{D}(\zeta) \cup \boldsymbol{D}(\zeta^p) \cup \cdots \cup \boldsymbol{D}(\zeta^{p^{r-1}})$. By (3-11) we conclude that the set $\mathrm{supp}(\Lambda_{D_0 p^{2m}}|_{\boldsymbol{D}(\zeta)})$ equals the set of preperiodic points of $\boldsymbol{t}$ in $\boldsymbol{D}(\zeta)$ of preperiod $m$. This completes the proof of (i).

To prove (ii), let $c$ in $]0, 1[$ be such that $\boldsymbol{B} \subseteq \boldsymbol{D}(z_0, c)$, let $\rho$ and $\kappa_c$ be given by Lemma 3.7 and let $M \geq 1$ be an integer such that $c\kappa_c^{rM} < \rho$. Let $D < 0$ be a discriminant and $z$ in $\mathrm{supp}(\Lambda_D) \cap \boldsymbol{B}$ be given. By (i) there is an integer $m \geq 0$ such that $\boldsymbol{t}^{rm}(z) = E_0$. Assume by contradiction that the least integer $m$ with this property satisfies $m > M$. Then by Lemma 3.7 and our choice of $M$ we have

$$|\boldsymbol{t}^{rM}(z) - E_0|_p \leq c\kappa_c^{rM} < \rho.$$

On the other hand, $\boldsymbol{t}^{r(m-M)}$ is injective on $\boldsymbol{D}(z_0, \rho)$ by Lemma 3.7(ii) and it maps $\boldsymbol{t}^{rM}(z)$ and $E_0$ to $E_0$, so $\boldsymbol{t}^{rM}(z) = E_0$. This contradicts the minimality of $m$ and proves that for every $z$ in $\mathrm{supp}(\Lambda_D) \cap \boldsymbol{B}$ we have $\boldsymbol{t}^{rM}(z) = E_0$. Equivalently,

$$\mathrm{supp}(\Lambda_D|_{\boldsymbol{B}}) \subseteq \bigcup_{i=1}^{M} \boldsymbol{t}^{-ir}(E_0).$$

Since this last set is finite and independent of $D$, this proves (ii) and completes the proof of the theorem. □

## 4. CM points in the supersingular reduction locus

The goal of this section is to prove the following result on the asymptotic distribution of CM points in the supersingular reduction locus. From this result and Theorem 3.5(ii), we deduce Theorem A at the end of this section.

**Theorem 4.1.** *For every $e$ in $Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)$ fix an arbitrary $\gamma_e$ in $\boldsymbol{D}(j(e))$ and for $r$ in $]0,1[$, put*

$$\boldsymbol{B}(r) := \bigcup_{e \in Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)} \boldsymbol{D}(\gamma_e, r).$$

*Then the following properties hold*:

(i) *For every $r$ in $]0,1[$ there exists $m > 0$ such that for every discriminant $D < 0$ satisfying $\mathrm{ord}_p(D) \geq m$, we have $\deg(\Lambda_D|_{\boldsymbol{B}(r)}) = 0$.*

(ii) *For every $m > 0$ there exists $r$ in $]0,1[$ such that for every $p$-supersingular discriminant $D < 0$ satisfying $\mathrm{ord}_p(D) \leq m$, we have $\mathrm{supp}(\Lambda_D) \subseteq \boldsymbol{B}(r)$.*

We present the proof of Theorem 4.1 in Section 4C below. In Section 4A we recall the definition of Katz' valuation. For that purpose, we briefly review Katz' theory of algebraic modular forms and the interpretation of the Eisenstein series $\mathrm{E}_{p-1}$ as an algebraic modular form over $\mathbb{Q} \cap \mathbb{Z}_p$. In Section 4B we use Katz–Lubin's extension of the theory of canonical subgroups to not too supersingular elliptic curves to give a description of the action of Hecke correspondences on the supersingular locus (Section 4B). For $p = 2$ and 3, we also rely on certain congruences satisfied by certain Eisenstein series (Proposition A.1 in Appendix A). This description is used in the proof of Theorem 4.1 and also in Section 5C on Hecke orbits in the supersingular locus.

**4A.** ***Katz' valuation.*** In this section we define Katz' valuation, which is based on Katz' theory of algebraic modular forms, and give an explicit formula relating it to the $j$-invariant (Proposition 4.3).

For the reader's convenience we start with a short review of Katz' theory of algebraic modular forms. For details see [Katz 1973, Chapter 1]. Let $k \in \mathbb{Z}$ be an integer and let $R_0$ be a ring (commutative and with identity). Denote by $R_0$-Alg the category of $R_0$-algebras. Given an $R_0$-algebra $R$, define an elliptic curve $E$ over $R$ as a proper, smooth morphism of schemes $E \to \mathrm{Spec}(R)$, whose geometric fibers are connected curves of genus one, together with a section $\mathrm{Spec}(R) \to E$, and denote by $\Omega^1_{E/R}$ the invertible sheaf of differential forms of degree 1 of $E$ over $R$. By replacing $\mathrm{Spec}(R)$ by an appropriate affine subset we can assume that $\Omega^1_{E/R}$ admits a nowhere vanishing global section. In this paper we assume, for simplicity, that this is always the case and denote by $\Omega^1_{E/R}(E)'$ the (nonempty) set of nowhere vanishing global sections of $\Omega^1_{E/R}$. An algebraic modular form $F$ of weight $k$ and level one over $R_0$ is a family of maps

$$F_R \colon \{(E, \omega) : E \text{ elliptic curve over } R, \omega \in \Omega^1_{E/R}(E)'\} \to R \quad (R \in R_0\text{-Alg}),$$

satisfying the following properties:

(i) $F_R(E, \omega)$ depends only on the isomorphism class of the pair $(E, \omega)$. More precisely, for every isomorphism of elliptic curves $\varphi \colon E \to E'$ over $R$, we have $F_R(E', \varphi_*\omega) = F_R(E, \omega)$. Here, $\varphi_*\omega$ denotes the push-forward of $\omega$ by $\varphi$.

(ii) $F_R(E, \lambda\omega) = \lambda^{-k} F_R(E, \omega)$ for every $\lambda \in R^\times$.

(iii) $F_R$ is compatible with base change. Namely, for every $R_0$-algebra morphism $g\colon R \to R'$, for the base change $(E,\omega)_{R'}$ of $(E,\omega)$ to $R'$ by $g$ we have $F_{R'}((E,\omega)_{R'}) = g(F_R(E,\omega))$.

Taking into account property (iii), from now on we simply write $F$ instead of $F_R$. Moreover, let $R_1$ be an $R_0$-algebra. Then, property (iii) ensures that $F$ induces an algebraic modular form $F_1$ over $R_1$. We say that $F_1$ is the *base change* of $F$ to $R_1$. We also say that $F$ is a *lifting* of $F_1$ to $R$.

Let $q$ be a formal variable and denote by $\mathrm{Tate}(q)$ the *Tate curve*, which is an elliptic curve over the field of fractions $\mathbb{Z}((q))$ of the ring of formal power series $\mathbb{Z}[\![q]\!]$; see [Katz 1973, Appendix 1]. The $j$-invariant of $\mathrm{Tate}(q)$ has the form

$$j(\mathrm{Tate}(q)) = \tfrac{1}{q} + 744 + \sum_{n=1}^{\infty} c_n q^n, \quad c_n \in \mathbb{Z}. \tag{4-1}$$

The *$q$-expansion* of an algebraic modular form $F$ over $R_0$ as above is defined as the element $F(q) \in \mathbb{Z}((q)) \otimes_{\mathbb{Z}} R_0$ obtained by evaluating $F$ at the pair $(\mathrm{Tate}(q), \omega_{\mathrm{can}})$ consisting of the Tate curve together with its canonical differential $\omega_{\mathrm{can}}$, both considered over $\mathbb{Z}((q)) \otimes_{\mathbb{Z}} R_0$. Moreover, $F$ is said to be *holomorphic at infinity* if $F(q) \in \mathbb{Z}[\![q]\!] \otimes_{\mathbb{Z}} R_0$.

Now, we state a version of the *$q$-expansion principle*, which is a particular case of [Katz 1973, Corollary 1.9.1].

**Theorem 4.2.** *Let $R_0$ be a ring and let $K \supseteq R_0$ be a $R_0$-algebra. Let $k \in \mathbb{Z}$ be an integer and let $F$ be an algebraic modular form over $K$ of weight $k$, level one and holomorphic at infinity. Assume that $F(q) \in \mathbb{Z}((q)) \otimes_{\mathbb{Z}} R_0$. Then, $F$ is the base change of a unique algebraic modular form over $R_0$ of weight $k$.*

There is a natural link between the previous theory and the classical theory of modular forms. We refer to [Katz 1973, Section A1.1] for details. For each classical holomorphic modular form of weight $k$ and level one $f\colon \mathbb{H} \to \mathbb{C}$, there exists a unique algebraic modular form $F$ over $\mathbb{C}$ associated to $f$ that is holomorphic at infinity. The Fourier expansion at infinity of $f$ and the $q$-expansion of $F$ are related by

$$f(\tau) = \sum_{n=0}^{\infty} a_n e^{2\pi i n\tau} \quad \text{if and only if} \quad F(q) = \sum_{n=0}^{\infty} a_n q^n.$$

For an even integer $k \geq 4$, let $\mathrm{E}_k$ be the *normalized Eisenstein series*

$$\mathrm{E}_k(\tau) = 1 - \frac{2k}{B_k} \sum_{n=1}^{\infty} \sigma_{k-1}(n) e^{2\pi i n\tau}, \quad \tau \in \mathbb{H}.$$

Here, the symbol $B_k$ denotes the $k$-th Bernoulli number and $\sigma_{k-1}(n) := \sum_{d\,|\,n, d>0} d^{k-1}$. The complex function $\mathrm{E}_k$ is a classical holomorphic modular form of weight $k$ and level one. Then, this function induces an algebraic modular form over $\mathbb{C}$, which we also denote by $\mathrm{E}_k$, having the $q$-expansion with rational coefficients

$$\mathrm{E}_k(q) = 1 - \frac{2k}{B_k} \sum_{n=1}^{\infty} \sigma_{k-1}(n) q^n. \tag{4-2}$$

When $p \geq 5$ and $k = p-1$, the von Staudt–Clausen theorem ensures that $\operatorname{ord}_p\left((2k)B_k^{-1}\right) = 1$. In particular, the coefficients of the Fourier expansion of $\mathrm{E}_{p-1}$ lie in $\mathbb{Z}_{(p)} := \mathbb{Q} \cap \mathbb{Z}_p$. Hence, by Theorem 4.2, we can consider $\mathrm{E}_{p-1}$ as an algebraic modular form of weight $p-1$ over $\mathbb{Z}_{(p)}$. On the other hand, the same reasoning and a direct examination of the Fourier expansions of $\mathrm{E}_4$ and $\mathrm{E}_6$ allow us to consider these Eisenstein series as algebraic modular forms of weight four and six over $\mathbb{Z}$.

For $E$ in $Y_{\mathrm{sups}}(\mathbb{C}_p)$, which we regard as an elliptic curve over $\mathcal{O}_p$, choose $\omega$ in $\Omega^1_{E/\mathcal{O}_p}(E)'$ and define *Katz' valuation*

$$v_p(E) := \begin{cases} \operatorname{ord}_p(\mathrm{E}_{p-1}(E,\omega)) & \text{if } p \geq 5; \\ \frac{1}{3} \cdot \operatorname{ord}_3(\mathrm{E}_6(E,\omega)) & \text{if } p = 3; \\ \frac{1}{4} \cdot \operatorname{ord}_2(\mathrm{E}_4(E,\omega)) & \text{if } p = 2. \end{cases}$$

Since for every $\lambda$ in $\mathcal{O}_p^\times$ we have $\mathrm{E}_k(E,\lambda\omega) = \lambda^{-k}\,\mathrm{E}_k(E,\omega)$, this definition does not depend on the particular choice of $\omega$. The above definition is motivated by the following considerations. The Hasse invariant $A_{p-1}$ is the unique algebraic modular form of weight $p-1$ over $\mathbb{F}_p$ with $q$-expansion $A_{p-1}(q) = 1$; see [Katz 1973, Chapter 2]. When $p \geq 5$, the base change to $\mathbb{F}_p$ of the form $\mathrm{E}_{p-1}$ equals $A_{p-1}$. On the other hand, when $p$ equals 2 or 3 it is not possible to lift $A_{p-1}$ to an algebraic modular form of level one, holomorphic at infinity, over $\mathbb{Z}_{(p)}$. However, the base change of $\mathrm{E}_4$ (resp. $\mathrm{E}_6$) to $\mathbb{F}_2$ (resp. to $\mathbb{F}_3$) is $A_1^4$ (resp. $A_2^3$). See Appendix A for details.

Since the Hasse invariant vanishes at supersingular elliptic curves, for every $E$ in $Y_{\mathrm{sups}}(\mathbb{C}_p)$ we have that $0 < v_p(E) \leq \infty$. An elliptic curve $E$ in $Y_{\mathrm{sups}}(\mathbb{C}_p)$ is *not too supersingular* if $v_p(E) < p/(p+1)$, and it is *too supersingular* otherwise.

The following result gives an explicit relation between $v_p(E)$ and $j(E)$. For $e$ in $Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)$, we use the number $\delta_e$ defined by (3-2) in Section 3A.

**Proposition 4.3.** *For each $e$ in $Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)$, denote by $\mathrm{j}_e$ the $j$-invariant of the unique zero of $\mathrm{E}_{p-1}$ (resp. $\mathrm{E}_4, \mathrm{E}_6$) in $\boldsymbol{D}(e)$ if $p \geq 5$ (resp. $p = 2, 3$). Then, for every $E$ in $Y_{\mathrm{sups}}(\mathbb{C}_p)$ we have*

$$v_p(E) = \sum_{e \in Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)} \frac{1}{\delta_e} \operatorname{ord}_p(j(E) - \mathrm{j}_e).$$

*Moreover, if $p \geq 5$ and $\mathrm{j}_e \equiv 0$ (resp. $\mathrm{j}_e \equiv 1728$) $\bmod \mathcal{M}_p$, then $\mathrm{j}_e = 0$ (resp. $\mathrm{j}_e = 1728$). In the case $p = 2$ (resp. $p = 3$), $Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)$ has a unique element $e$ and $\mathrm{j}_e = 0$ (resp. $\mathrm{j}_e = 1728$).*

It follows from the proof of this proposition that for every $e$ in $Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)$ the number $\mathrm{j}_e$ is algebraic over $\mathbb{Q}$ and is in the quadratic unramified extension of $\mathbb{Q}_p$. We note that in the case $\mathrm{j}_e \not\equiv 0, 1728 \bmod \mathcal{M}_p$, the elliptic curve class whose $j$-invariant is $\mathrm{j}_e$ is not CM,[§] but it is "fake CM" in the sense of [Coleman and McMurdy 2006]; see Remark 4.4 below.

[§] In fact, $\mathrm{j}_e$ need not be an algebraic integer: For $p = 13$ (resp. 17, 19, 23) there is a unique $e$ in $Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)$ whose $j$-invariant is different from 0 and 1728, and we have $\mathrm{j}_e = 2^7 \cdot 3^3 \cdot 5^3/691$ (resp. $2^{10} \cdot 3^3 \cdot 5^3/3617$, $2^8 \cdot 3^3 \cdot 5^3 \cdot 11/43867$, $2^8 \cdot 3^3 \cdot 5^3 \cdot 41/(131 \cdot 593)$).

*Proof of Proposition 4.3.* Assume $p \geq 5$, so $p - 1 \not\equiv 2, 8 \mod 12$. We can thus write $p-1$ uniquely in the form $p - 1 = 12m + 4\delta + 6\varepsilon$ with $m \geq 0$ integer and $\delta, \varepsilon \in \{0, 1\}$. The modular discriminant

$$\Delta(\tau) = e^{2\pi i\tau} \prod_{n=1}^{\infty} (1 - e^{2\pi in\tau})^{24}, \quad \tau \in \mathbb{H},$$

is a classical holomorphic modular form of weight 12 and level one; see, e.g., [Diamond and Shurman 2005, Sections 1.1 and 1.2]. The infinite product above shows that the Fourier coefficients of $\Delta$ are rational integers. Hence, Theorem 4.2 ensures that $\Delta$ can be considered as an algebraic modular form over $\mathbb{Z}$. At the level of classical modular forms, we have the identity

$$\mathrm{E}_{p-1} = \Delta^m \, \mathrm{E}_4^{\delta} \, \mathrm{E}_6^{\varepsilon} \, P(j),$$

where $P(X)$ is a monic polynomial over $\mathbb{Z}_{(p)}$ of degree $m$ such that $P_{\mathrm{sups}}(X) := X^{\delta}(X - 1728)^{\varepsilon} P(X)$ reduces modulo $p$ to the supersingular polynomial, i.e., the monic separable polynomial over $\mathbb{F}_p$ whose roots are the $j$-invariants of the supersingular elliptic curves over $\bar{\mathbb{F}}_p$; see, e.g., [Kaneko and Zagier 1998, Theorem 1]. Using the classical identities $\mathrm{E}_4^3 = \Delta j$ and $\mathrm{E}_6^2 = \Delta(j - 1728)$ we get

$$\mathrm{E}_{p-1}^{12} = \Delta^{p-1} j^{4\delta} (j - 1728)^{6\varepsilon} P(j)^{12}.$$

Theorem 4.2 ensures that the above identity also holds at the level of algebraic modular forms over $\mathbb{Z}_{(p)}$. Write

$$P_{\mathrm{sups}}(X) = \prod_{e \in Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)} (X - \mathrm{j}_e),$$

where $\mathrm{j}_e \in \boldsymbol{D}\big(j(e)\big)$ for each $e \in Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)$. Now, for every pair $(E, \omega)$ over $\mathcal{O}_p$ having good reduction we have $\Delta(E, \omega) \in \mathcal{O}_p^{\times}$, hence

$$|\mathrm{E}_{p-1}(E, \omega)|_p^{12} = |j(E)|_p^{4\delta} |j(E) - 1728|_p^{6\varepsilon} \prod_{\substack{e \in Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p) \\ \mathrm{j}_e \not\equiv 0, 1728}} |j(E) - \mathrm{j}_e|_p^{12}.$$

Since $p \geq 5$, we have that $j = 0$ (resp. $j = 1728$) is supersingular at $p$ if and only if $p \equiv 2 \mod 3$ (resp. $p \equiv 3 \mod 4$) [Silverman 2009, Chapter V, Examples 4.4 and 4.5]. This implies the result when $p \geq 5$. The cases $p = 2$ and 3 follow similarly from the formulas

$$|\mathrm{E}_4(E, \omega)|_2^3 = |j(E)|_2 \quad \text{and} \quad |\mathrm{E}_6(E, \omega)|_3^2 = |j - 1728|_3,$$

respectively. This completes the proof of the proposition. □

**Remark 4.4.** Let $e$ in $Y_{\mathrm{sups}}(\mathbb{C}_p)$ be such that $\mathrm{j}_e \not\equiv 0, 1728 \mod \mathcal{M}_p$, and let $E_e$ be the elliptic curve class in $Y(\mathbb{C}_p)$ such that $j(E_e) = \mathrm{j}_e$. Then $E_e$ is not CM, but it is "fake CM" in the sense of [Coleman and McMurdy 2006]. In particular, $\mathrm{j}_e$ is not a singular modulus over $\mathbb{C}_p$. To show that $E_e$ is not CM, choose a field isomorphism $\mathbb{C}_p \simeq \mathbb{C}$ and $\tau_e$ in $\mathbb{H}$ such that $E_e(\mathbb{C}) \simeq \mathbb{C}/(\mathbb{Z} + \tau_e\mathbb{Z})$. It is sufficient to show that $\tau_e$ is transcendental over $\mathbb{Q}$; see, e.g., [Lang 1973, Chapter 1, Section 5]. The complex number $\tau_e$ must be a

zero of the holomorphic function $\tau \mapsto \mathrm{E}_{p-1}(\tau)$. Since $j(\tau_e) = \mathrm{j}_e$ is different from 0 and 1728, it follows that $\tau_e$ is not equivalent to $\rho = \frac{1}{2}(1+\sqrt{-3})$ or $i = \sqrt{-1}$ under the action of the modular group $\mathrm{SL}_2(\mathbb{Z})$ by Möbius transformations on $\mathbb{H}$. Then [Kohnen 2003, Theorem 1] implies that $\tau_e$ is transcendental over $\mathbb{Q}$.

To see that $E_e$ is fake CM, note first that, since the reduction modulo $p$ of $P_{\mathrm{sups}}(X)$ is separable and splits completely over $\mathbb{F}_{p^2}$, by Hensel's lemma all roots of $P_{\mathrm{sups}}(X)$ are in the ring of integers $\mathcal{O}$ of the unramified quadratic extension of $\mathbb{Q}_p$. As $\mathrm{j}_e$ is a root of $P_{\mathrm{sups}}(X)$, this implies that $E_e$ represents an elliptic curve over $\mathcal{O}$. Let $[p]_e$ and $\phi$ be the multiplication by $p$ and the $p^2$-power Frobenius endomorphism on the supersingular curve $e$, respectively. Then there exists $\sigma$ in $\mathrm{Aut}(e)$ satisfying $\sigma \circ [p]_e = \phi$; see [Silverman 2009, Chapter II, Corollary 2.12]. Since $j(e) = \pi(\mathrm{j}_e)$ is different from 0 and 1728, we have $\sigma = \pm 1$ and $\pm[p]_e = \phi$. Choose $\pi_0 = \pm p$ as a uniformizer of $\mathcal{O}$. The multiplication by $\pi_0$ map on the formal group $\mathcal{F}_{E_e}$ of $E_e$ defines an endomorphism $f(X)$ of $\mathcal{F}_{E_e}$, satisfying

$$f(X) \equiv \pi_0 X \bmod X^2 \quad \text{and} \quad f(X) \equiv X^{p^2} \bmod \pi_0.$$

It follows that $\mathcal{F}_{E_e}$ is a Lubin–Tate formal group over $\mathcal{O}$; see [Hazewinkel 1978, Section 8], and compare with [Coleman and McMurdy 2006, Remark 3.4]. In particular $\mathrm{End}(\mathcal{F}_{E_e}) \simeq \mathcal{O}$ and therefore $E_e$ is fake CM; see [Hazewinkel 1978, Theorem 8.1.5 and Proposition 23.2.6].

**4B.** ***Katz' kite.*** The goal of this section is to give the following description of the action of Hecke correspondences on the supersingular locus.

**Proposition 4.5.** *Let* $\hat{v}_p \colon Y_{\mathrm{sups}}(\mathbb{C}_p) \to [0,\, p/(p+1)]$ *be the map defined by*

$$\hat{v}_p := \min\left\{v_p,\, \frac{p}{p+1}\right\}.$$

*Moreover, denote by* $\tau_0$ *the identity on* $\mathrm{Div}([0,\, p/(p+1)])$, *let* $\tau_1$ *be the piecewise-affine correspondence on* $[0,\, p/(p+1)]$ *defined by*

$$\tau_1(x) := \begin{cases} [px] + p[x/p] & \text{if } x \in [0,\, 1/(p+1)]; \\ [1-x] + p[x/p] & \text{if } x \in \,]1/(p+1),\, p/(p+1)], \end{cases}$$

*and for each integer* $m \geq 2$ *define the correspondence* $\tau_m$ *on* $[0,\, p/(p+1)]$ *recursively, by*

$$\tau_m := \tau_1 \circ \tau_{m-1} - p\tau_{m-2}.$$

*Then for every integer* $m \geq 0$ *and every integer* $n_0 \geq 1$ *not divisible by* $p$, *we have*

$$(\hat{v}_p)_* \circ T_{p^m n_0}|_{Y_{\mathrm{sups}}(\mathbb{C}_p)} = \sigma_1(n_0) \cdot \tau_m \circ (\hat{v}_p)_*.$$

See Figure 1 for the graph of the correspondence $\tau_1$ and Lemma 5.7 in Section 5C for a formula of $\tau_m$ for every $m \geq 0$.

The proof of Proposition 4.5 is given after a couple of lemmas. The following is a reformulation, in our setting, of a theorem of Katz–Lubin on the existence of canonical subgroups for elliptic curves that are not too supersingular; see [Katz 1973, Theorems 3.1 and 3.10.7] and also [Buzzard 2003, Theorem 3.3].

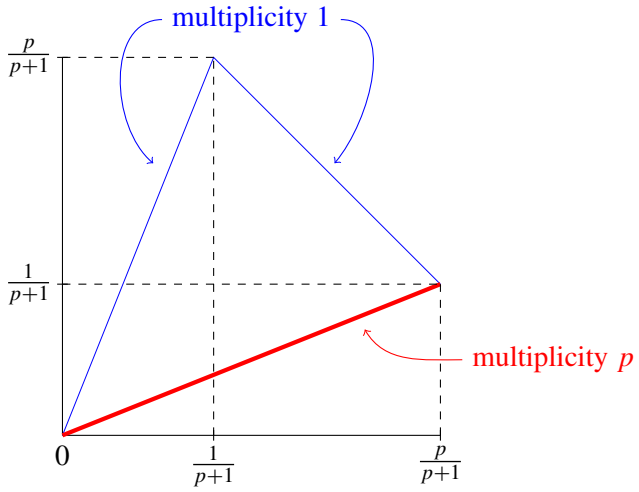



**Figure 1.** Graph of the correspondence $\tau_1$ representing the action of $T_p$ in terms of the projection $\hat{v}_p$.

**Lemma 4.6.** *For every elliptic curve $E$ in $Y_{\mathrm{sups}}(\mathbb{C}_p)$ that is not too supersingular there is a unique subgroup $H(E)$ of $E$ of order $p$ satisfying*

$$\hat{v}_p(E/H(E)) = \begin{cases} p v_p(E) & \text{if } v_p(E) \in ]0, 1/(p+1)]; \\ 1 - v_p(E) & \text{if } v_p(E) \in ]1/(p+1), p/(p+1)[. \end{cases} \tag{4-3}$$

*Furthermore, $H(E)$ is also uniquely characterized by the property that for every subgroup $C$ of $E$ of order $p$ that is different from $H(E)$, we have*

$$v_p(E/C) = p^{-1} v_p(E). \tag{4-4}$$

*In addition, the map*

$$\begin{aligned} \boldsymbol{t} \colon \left\{ E \in Y_{\mathrm{sups}}(\mathbb{C}_p) : v_p(E) < \tfrac{p}{p+1} \right\} &\to Y_{\mathrm{sups}}(\mathbb{C}_p) \\ E &\mapsto \boldsymbol{t}(E) := E/H(E) \end{aligned}$$

*satisfies the following properties*:

(i) *Let $E$ be in $Y_{\mathrm{sups}}(\mathbb{C}_p)$ and let $C$ be a subgroup of $E$ of order $p$. In the case $v_p(E) < p/(p+1)$, assume in addition that $C \neq H(E)$. Then*

$$v_p(E/C) = p^{-1} \hat{v}_p(E) \quad \text{and} \quad \boldsymbol{t}(E/C) = E.$$

(ii) *For $E$ in $Y_{\mathrm{sups}}(\mathbb{C}_p)$ satisfying $1/(p+1) < v_p(E) < p/(p+1)$, we have $\boldsymbol{t}^2(E) = E$.*

*Proof.* For $E$ in $Y_{\mathrm{sups}}(\mathbb{C}_p)$ that is not too supersingular, note that the uniqueness statements about $H(E)$ follow from the fact that (4-3) and (4-4) imply that $H(E)$ is the unique subgroup $C$ of $E$ of order $p$ satisfying $v_p(E/C) \neq p^{-1} v_p(E)$.

Assume $p \geq 5$ and let $E$ be an elliptic curve in $Y_{\mathrm{sups}}(\mathbb{C}_p)$ that is not too supersingular, so that $v_p(E) < p/(p+1)$. Let $\omega$ be a differential form in $\Omega^1_{E/\mathcal{O}_p}(E)'$ and put $r_E := \mathrm{E}_{p-1}(E,\omega) \in \mathcal{O}_p$. Since $\overline{\mathbb{C}_p^{\mathrm{unr}}}$ and $\mathbb{C}_p$ have the same valuation group we can find $r \in \overline{\mathbb{C}_p^{\mathrm{unr}}}$ satisfying $\mathrm{ord}_p(r) = \mathrm{ord}_p(r_E)$. Then $r$ lies in the ring of integers $R_0$ of some finite extension of $\mathbb{C}_p^{\mathrm{unr}}$, and $R_0$ is a complete discrete valuation ring of residue characteristic $p$ and generic characteristic zero. The triple $(E, \omega, rr_E^{-1})$ defines a $r$-situation in the sense of [Katz 1973, Theorem 3.1] (see also [loc. cit., Section 2.2]) and therefore there is a canonical subgroup $H(E)$ of $E$ of order $p$. Then [loc. cit., Theorem 3.10.7(2, 3)] implies (4-3) and (ii), see also the proof of [Buzzard 2003, Theorem 3.3(iii)], and (4-4) and (i) are given by [Katz 1973, Theorem 3.10.7(5)]. Finally, note that for $E$ in $Y_{\mathrm{sups}}(\mathbb{C}_p)$ satisfying $v_p(E) \geq p/(p+1)$, the assertion (i) follows from [loc. cit., Theorem 3.10.7(4)]. This completes the proof of the proposition in the case $p \geq 5$.

It remains to prove the proposition in the cases $p = 2$ and $p = 3$. We only give the proof in the case $p = 2$, the case $p = 3$ being analogous. Let $\mathrm{E}_1$ be an algebraic modular form of weight one and level $n_1$, with $3 \leq n_1 \leq 11$ odd, holomorphic at infinity and defined over $\mathbb{Z}[1/n_1]$ whose reduction modulo 2 is $A_1$; see Appendix A for details on level structures. Let $E$ in $Y_{\mathrm{sups}}(\mathbb{C}_2)$ be an elliptic curve that is not too supersingular, let $\omega$ be a differential form in $\Omega^1_{E/\mathcal{O}_2}(E)'$ and $\alpha_{n_1}$ a level $n_1$ structure on $E$ over $\mathcal{O}_2$. By Proposition A.1 and our hypothesis $v_2(E) < \frac{2}{3}$, we have

$$\mathrm{ord}_2(\mathrm{E}_1(E, \omega, \alpha_{n_1})) = v_2(E) < \tfrac{2}{3}.$$

Then, [Katz 1973, Theorem 3.1] gives the existence of $H(E)$ which might depend on the choice of $\alpha_{n_1}$. The fact that $H(E)$ depends only on $E$ follows from the characterization in [loc. cit., Theorem 3.10.7(1)] of the canonical subgroup as the subgroup of order 2 containing the unique point corresponding to the solution with valuation $1 - v_2(E)$ of the equation $[2](X) = 0$ in the formal group of $E$ (here $[2]$ denotes the multiplication by 2 map and $X$ is a certain normalized parameter for the formal group). Then (4-3), (4-4), (i) and (ii) follow from [loc. cit., Theorem 3.10.7] as in the case $p \geq 5$ above. This completes the proof of the lemma. □

**Lemma 4.7.** *Let $E$ in $Y_{\mathrm{sups}}(\mathbb{C}_p)$ be such that*

$$v_p(E) < \begin{cases} 1 & \text{if } p \geq 5; \\ (2p-1)/(2p) & \text{if } p = 2 \text{ or } 3. \end{cases}$$

*Then for every subgroup $C$ of $E$ of order not divisible by $p$, we have $v_p(E/C) = v_p(E)$.*

*Proof.* For $E_0$ in $Y(\mathbb{C}_p)$ and $\zeta$ in $\mathbb{Z}_p$, denote by $[\zeta]_{E_0}$ the multiplication by $\zeta$ map in the formal group of $E_0$.

Put $E' := E/C$ and denote by $\phi\colon E \to E'$ an isogeny with kernel $C$. Let $X$ (resp. $Y$) be a parameter of the formal group of $E$ (resp. $E'$), such that for any $(p-1)$-th root of unity $\zeta \in \mathbb{Z}_p$ we have $[\zeta]_E(X) = \zeta X$ (resp. $[\zeta]_{E'}(Y) = \zeta Y$); see [Katz 1973, Lemma 3.6.2(2)]. Let $\omega$ be a differential form in $\Omega^1_{E/\mathcal{O}_p}(E)'$ whose expansion in the parameter $X$ is of the form

$$\omega = \left(1 + \sum_{n=1}^{\infty} a_n X^n\right) dX,$$

where $a_n \in \mathcal{O}_p$ for all $n \geq 1$. Then, by [loc. cit., Proposition 3.6.6] we have

$$[p]_E(X) = pX + aX^p + \sum_{m\geq 2} c_m X^{m(p-1)+1},$$

where $c_m \in \mathcal{O}_p$ for all $m \geq 2$ and $a \in \mathcal{O}_p$ satisfies

$$a \equiv A_{p-1}((E,\omega)_{\mathcal{O}_p/p\mathcal{O}_p}) \bmod p\mathcal{O}_p, \tag{4-5}$$

where $(E,\omega)_{\mathcal{O}_p/p\mathcal{O}_p}$ denotes the base change of $(E,\omega)$ to $\mathcal{O}_p/p\mathcal{O}_p$. Similarly,

$$[p]_{E'}(Y) = pY + a'Y^p + \sum_{m\geq 2} c'_m Y^{m(p-1)+1},$$

where $c'_m \in \mathcal{O}_p$ for all $m \geq 2$ and $a' \in \mathcal{O}_p$ satisfies, for some differential form $\omega'$ of $\Omega^1_{E'/\mathcal{O}_p}(E)'$,

$$a' \equiv A_{p-1}((E',\omega')_{\mathcal{O}_p/p\mathcal{O}_p}) \bmod p\mathcal{O}_p. \tag{4-6}$$

Since the order of $\mathrm{Ker}(\phi) = C$ is not divisible by $p$, the isogeny $\phi$ induces an isomorphism of formal groups of the form

$$\phi(X) = \sum_{n=1}^{\infty} t_n X^n,$$

where $t_n \in \mathcal{O}_p$ for all $n \geq 1$. Since $\phi(X)$ is invertible, we must have $t_1 \in \mathcal{O}_p^\times$. By the identity $[p]_{E'} \circ \phi = \phi \circ [p]_E$ we get

$$p(t_1X + t_2X^2 + t_3X^3 + \cdots) + a'(t_1X + t_2X^2 + t_3X^3 + \cdots)^p + \cdots \\ = t_1(pX + aX^p + \cdots) + t_2(pX + aX^p + \cdots)^2 + \cdots$$

Comparing the coefficients of $X^p$, we get

$$pt_p + a't_1^p = t_1 a + t_p p^p.$$

Using that $t_1 \in \mathcal{O}_p^\times$ we obtain

$$\mathrm{ord}_p(a') = \mathrm{ord}_p(a't_1^{p-1}) = \mathrm{ord}_p(a + t_1^{-1}t_p(p^p - p)). \tag{4-7}$$

In the case $p \geq 5$, (4-5) implies $\mathrm{ord}_p(a - \mathrm{E}_{p-1}(E,\omega)) \geq 1$, so by our hypothesis $v_p(E) < 1$ we have $\mathrm{ord}_p(a) = v_p(E) < 1$. Combined with (4-7), this implies $\mathrm{ord}_p(a') = \mathrm{ord}_p(a) = v_p(E) < 1$. Finally, by (4-6) we have $\mathrm{ord}_p(a' - \mathrm{E}_{p-1}(E',\omega')) \geq 1$, so $v_p(E') = \mathrm{ord}_p(a') = v_p(E)$. This proves the lemma in the case $p \geq 5$. For the case $p = 2$ or 3, (4-5), (4-6), (4-7), our hypothesis $v_p(E) < (2p-1)/(2p)$ and Proposition A.1 imply in a similar way

$$\mathrm{ord}_p(a) = v_p(E) < \frac{2p-1}{2p}, \quad \mathrm{ord}_p(a') = \mathrm{ord}_p(a) \quad \text{and} \quad v_p(E') = \mathrm{ord}_p(a').$$

This completes the proof of the lemma. □

*Proof of Proposition 4.5.* By the multiplicative property of Hecke correspondences (2-6) and Lemma 4.7, it is sufficient to consider the case $n_0 = 1$. Moreover, in view of (2-5) and the recursive definition of $\tau_m$ for $m \geq 2$, it is sufficient to consider the case $m = 1$. For $E$ in $Y_{\mathrm{sups}}(\mathbb{C}_p)$ satisfying $\hat{v}_p(E) < p/(p+1)$, this is given by (4-3) and (4-4) in Lemma 4.6, together with the fact that $\deg(T_p(E)) = p+1$. Finally, for $E$ in $Y_{\mathrm{sups}}(\mathbb{C}_p)$ satisfying $\hat{v}_p(E) = p/(p+1)$ the desired statement follows from Lemma 4.6(i). This completes the proof of the proposition. □

**4C.** ***Proof of Theorem 4.1.*** The proof of Theorem 4.1 is below, after a couple of lemmas.

**Lemma 4.8.** *Let $D < 0$ be a discriminant and let $E$ and $E'$ be in* $\mathrm{supp}(\Lambda_D)$. *Then, for every integer $m \geq 1$ there exists an isogeny $E \to E'$ of degree coprime to $m$.*

*Proof.* Denote by $d$ and $f$ the fundamental discriminant and conductor of $D$, respectively, and fix a field isomorphism $\mathbb{C}_p \simeq \mathbb{C}$. Since $E$ and $E'$ are CM with ring of endomorphisms isomorphic to $\mathcal{O}_{d,f}$, we can find proper fractional $\mathcal{O}_{d,f}$-ideals $\mathfrak{a}$ and $\mathfrak{a}'$ in $\mathbb{Q}(\sqrt{D})$ for which we have the complex uniformizations $E(\mathbb{C}) \simeq \mathbb{C}/\mathfrak{a}$ and $E'(\mathbb{C}) \simeq \mathbb{C}/\mathfrak{a}'$. Then there is a natural identification

$$\iota\colon \mathrm{Hom}(E, E') \to \mathfrak{a}'\mathfrak{a}^{-1} = \{\lambda \in \mathbb{C} : \lambda\mathfrak{a} \subseteq \mathfrak{a}'\}.$$

Without loss of generality, assume $\mathfrak{a}' \subset \mathfrak{a}$, and choose $\mathbb{Z}$-generators $\alpha$ and $\beta$ of the ideal $\mathfrak{a}'\mathfrak{a}^{-1}$ of $\mathcal{O}_{d,f}$. Then

$$f(x, y) := (\alpha x - \beta y)\overline{(\alpha x - \beta y)}/[\mathcal{O}_{d,f} : \mathfrak{a}'\mathfrak{a}^{-1}]$$

is a positive definite primitive binary quadratic form with integer coefficients and discriminant $d$ [Cox 2013, Theorem 7.7 and Exercise 7.17]. Moreover, there are integers $x_0$ and $y_0$ such that $f(x_0, y_0)$ is coprime to $m$ [loc. cit., Lemma 2.25]. If we denote by $\phi_0$ the isogeny in $\mathrm{Hom}(E, E')$ satisfying $\lambda_0 := \iota(\phi_0) = \alpha x_0 - \beta y_0$, then

$$\begin{aligned}
\deg(\phi_0) &= \#\,\mathrm{Ker}(\phi_0) \\
&= [\mathfrak{a}' : \lambda_0\mathfrak{a}] \\
&= [\mathfrak{a}'\mathfrak{a}^{-1} : \lambda_0\mathcal{O}_{d,f}] \\
&= [\mathcal{O}_{d,f} : \lambda_0\mathcal{O}_{d,f}]/[\mathcal{O}_{d,f} : \mathfrak{a}'\mathfrak{a}^{-1}] \\
&= \lambda_0\overline{\lambda_0}/[\mathcal{O}_{d,f} : \mathfrak{a}'\mathfrak{a}^{-1}] \\
&= f(x_0, y_0).
\end{aligned}$$

This proves that $\deg(\phi_0)$ is coprime to $m$, and completes the proof of the lemma. □

The following lemma is analogous to [Coleman and McMurdy 2006, Lemma 4.8], which concerns $p \geq 3$ in the context of certain modular curves of level bigger than 1. See also [Gross 1986, Proposition 5.3].

**Lemma 4.9.** *Let $D$ be a $p$-supersingular discriminant and $m \geq 0$ the largest integer such that $p^m$ divides the conductor of $D$. Then for every $E$ in* $\mathrm{supp}(\Lambda_D)$ *we have*

$$\hat{v}_p(E) = \begin{cases} \frac{1}{2} \cdot p^{-m} & \text{if } p \text{ ramifies in } \mathbb{Q}(\sqrt{D}); \\ p/(p+1) \cdot p^{-m} & \text{if } p \text{ is inert in } \mathbb{Q}(\sqrt{D}).^{\P} \end{cases}$$

*Proof.* Let $d$ be the fundamental discriminant of $D$ and $f \geq 1$ the integer such that the conductor of $D$ is equal to $p^m f$, so $D = d(fp^m)^2$ and $f$ is not divisible by $p$. By Lemmas 4.7 and 4.8 with $m = p$, we deduce that for $E$ in $\mathrm{supp}(\Lambda_D)$ the number $\hat{v}_p(D) := \hat{v}_p(E)$ is independent of $E$. By Zhang's formula (2-9) with $\tilde{f} = p^m$ it follows that there exists an isogeny of degree $f$ from some elliptic curve in $\mathrm{supp}(\Lambda_{dp^{2m}})$ to an elliptic curve in $\mathrm{supp}(\Lambda_D)$. We conclude from Lemma 4.7 that $\hat{v}_p(D) = \hat{v}_p(dp^{2m})$. Thus, it is enough to prove the lemma in the case where $f = 1$.

We start with $m = 0$ and $m = 1$. By (2-11) with $f = 1$ and Proposition 4.5 with $m = 1$ and $n_0 = 1$, we have

$$\mathrm{supp}(\tau_1(\hat{v}_p(d))) = \begin{cases} \{\hat{v}_p(d), \hat{v}_p(dp^2)\} & \text{if } p \text{ ramifies in } \mathbb{Q}(\sqrt{d}); \\ \{\hat{v}_p(dp^2)\} & \text{if } p \text{ is inert in } \mathbb{Q}(\sqrt{d}). \end{cases}$$

From the definition of $\tau_1$ we have that $p/(p+1)$ is the only value of $x$ in $]0,\, p/(p+1)]$ such that $\tau_1(x)$ is supported on a single point. We conclude that if $p$ is inert in $\mathbb{Q}(\sqrt{d})$, then $\hat{v}_p(d) = p/(p+1)$ and therefore $\hat{v}_p(dp^2) = 1/(p+1)$. On the other hand, $\frac{1}{2}$ is the only value of $x$ in $]0,\, p/(p+1)]$ satisfying $x \in \mathrm{supp}(\tau_1(x))$. So, if $p$ ramifies in $\mathbb{Q}(\sqrt{d})$, then $\hat{v}_p(d) = \frac{1}{2}$ and therefore $\hat{v}_p(dp^2) = \frac{1}{2}p^{-1}$. This completes the proof of the lemma when $m = 0$ and $m = 1$. Assume $m \geq 2$ and note that by (2-12) with $f = 1$ and by Proposition 4.5 with $n_0 = 1$,

$$\{\hat{v}_p(dp^{2m})\} = \begin{cases} \mathrm{supp}\big((\tau_m - \tau_{m-1})\big(\frac{1}{2}\big)\big) & \text{if } p \text{ ramifies in } \mathbb{Q}(\sqrt{d}); \\ \mathrm{supp}((\tau_m - \tau_{m-2})(p/(p+1))) & \text{if } p \text{ is inert in } \mathbb{Q}(\sqrt{d}). \end{cases}$$

From the definition of $\tau_m$, we see that the right-hand side contains $\frac{1}{2} \cdot p^{-m}$ if $p$ ramifies in $\mathbb{Q}(\sqrt{d})$ and $p/(p+1) \cdot p^{-m}$ if $p$ is inert in $\mathbb{Q}(\sqrt{d})$. This proves $\hat{v}_p(dp^{2m}) = \frac{1}{2} \cdot p^{-m}$ in the former case and $\hat{v}_p(dp^{2m}) = p/(p+1) \cdot p^{-m}$ in the latter, and completes the proof of the lemma. □

*Proof of Theorem 4.1.* To prove (i), note that by Proposition 4.3 there is $m > 0$ so that $\hat{v}_p(\boldsymbol{B}(r)) \subseteq ]p/(p+1) \cdot p^{-m},\, p/(p+1)]$. Then by Lemma 4.9 for every $p$-supersingular discriminant $D < 0$ satisfying $\mathrm{ord}_p(D) \geq 2m + 3$ we have $\mathrm{supp}((\hat{v}_p)_*(\Lambda_D)) \cap \hat{v}_p(\boldsymbol{B}(r)) = \varnothing$, and therefore $\deg(\Lambda_D|_{\boldsymbol{B}(r)}) = 0$. On the other hand, if $D$ is a $p$-ordinary discriminant, then $\mathrm{supp}(\Lambda_D) \subset Y_{\mathrm{ord}}(\mathbb{C}_p)$ is disjoint from $\boldsymbol{B}(r)$, and therefore $\deg(\Lambda_D|_{\boldsymbol{B}(r)}) = 0$. This completes the proof of (i).

To prove (ii), note that by Proposition 4.3 there is $r$ in $]0, 1[$ so that

$$\hat{v}_p^{-1}\big(\big[\tfrac{1}{2} \cdot p^{-m},\, \tfrac{p}{p+1}\big]\big) \subseteq \boldsymbol{B}(r).$$

¶When $D = -3$ (resp. $D = -4$) is $p$-supersingular we have $j(E) = 1728$ (resp. 0) and $v_p(E) = \infty$, so in this formula we cannot replace the projection $\hat{v}_p$ by the valuation $v_p$. Compare with [Coleman and McMurdy 2006, Lemma 4.8].

Then by Lemma 4.9 for every $p$-supersingular discriminant $D < 0$ satisfying $\mathrm{ord}_p(D) \leq m$ we have $\mathrm{supp}((\hat{v}_p)_*(\Lambda_D)) \subseteq \left[\frac{1}{2} \cdot p^{-m},\, p/(p+1)\right]$ and therefore $\mathrm{supp}(\Lambda_D) \subseteq \boldsymbol{B}(r)$. This completes the proof of (ii) and of the theorem. □

*Proof of Theorem A.* In the case where all the discriminants in the sequence $(D_n)_{n=1}^{\infty}$ are $p$-ordinary (resp. $p$-supersingular), Theorem A is a direct consequence of Theorem 3.5(ii) (resp. Theorem 4.1), together with (2-8) and Lemma 2.3. The general case follows from these two special cases. □

## 5. Hecke orbits

The goal of this section is to prove Theorem C on the asymptotic distribution of Hecke orbits. The proof is divided into three complementary cases, according to whether the starting elliptic curve class has bad, ordinary or supersingular reduction. These are stated as Propositions 5.1, 5.2 and 5.6 in Sections 5A, 5B and 5C, respectively. In each case we prove a stronger quantitative statement.

**5A.** ***Hecke orbits in the bad reduction locus.*** In this section we prove a stronger version of the part of Theorem C concerning the bad reduction locus, which is stated as Proposition 5.1 below. We start by recalling some well-known results on the uniformization of $p$-adic elliptic curves with multiplicative reduction. See [Tate 1995] for the case of elliptic curves over complete discrete valued field, and [Roquette 1970] for the case of complete valued fields (see also [Silverman 1994, Chapter V, Theorem 3.1 and Remark 3.1.2]).

Let $z$ be in $\boldsymbol{D}(0,1)^* := \{z' \in \mathbb{C}_p : 0 < |z'|_p < 1\}$. We obtain, by the specialization $q = z$ in the Tate curve, an elliptic curve $\mathrm{Tate}(z)$ over $\mathbb{C}_p$ whose $j$-invariant satisfies

$$|j(\mathrm{Tate}(z))|_p = |z|_p^{-1} > 1, \tag{5-1}$$

see (4-1). This defines a bijective map

$$\begin{aligned} \boldsymbol{D}(0,1)^* &\to Y_{\mathrm{bad}}(\mathbb{C}_p) \\ z &\mapsto \mathrm{Tate}(z). \end{aligned}$$

Moreover, for each $z \in \boldsymbol{D}(0,1)^*$ there exists an explicit uniformization by $\mathbb{C}_p^\times$ of the set of $\mathbb{C}_p$-points of $\mathrm{Tate}(z)$. This uniformization induces an isomorphism of analytic groups $\varphi_z \colon \mathbb{C}_p^\times/z^{\mathbb{Z}} \to \mathrm{Tate}(z)(\mathbb{C}_p)$, see [Tate 1995, Theorem 1] for details. This allows us to give, for each integer $n \geq 1$, the following description of $T_n(\mathrm{Tate}(z))$. Note that for each positive divisor $k$ of $n$ and each $\ell \in \boldsymbol{D}(0,1)^*$ satisfying $\ell^k = z^{n/k}$, the set

$$C_{n,\ell} := \{a \in \mathbb{C}_p^\times : a^{n/k} \in \ell^{\mathbb{Z}}\}/z^{\mathbb{Z}} \tag{5-2}$$

is a subgroup of order $n$ of $\mathbb{C}_p^\times/z^{\mathbb{Z}}$. It is the kernel of the morphism of analytic groups $\mathbb{C}_p^\times/z^{\mathbb{Z}} \to \mathbb{C}_p^\times/\ell^{\mathbb{Z}}$ induced by the map $a \mapsto a^{n/k}$. Precomposing this morphism with $\varphi_z^{-1}$ and then composing with $\varphi_\ell$, we obtain an isogeny $\mathrm{Tate}(z) \to \mathrm{Tate}(\ell)$ of degree $n$ whose kernel is $\varphi_z(C_{n,\ell})$. Since every subgroup of order

$n$ of $\mathbb{C}_p^\times/z^{\mathbb{Z}}$ is of the form (5-2), we deduce that

$$T_n(\mathrm{Tate}(z)) = \sum_{\substack{k>0,\, k\,|\,n \\ \ell^k = z^{n/k}}} \mathrm{Tate}(\ell). \tag{5-3}$$

In the case where $E$ is in $Y_{\mathrm{bad}}(\mathbb{C}_p)$, Theorem C is a direct consequence of the following result together with (2-1), (2-2) and Lemma 2.3.

**Proposition 5.1.** *Let $z$ in $\boldsymbol{D}(0,1)^*$ and $R>1$ be given. Then, for every $\varepsilon>0$ there exists $C>0$ such that for every integer $n\geq 1$ we have*

$$\deg(T_n(\mathrm{Tate}(z))|_{\boldsymbol{D}^\infty(0,R)}) \leq Cn^{1/2}d(n).$$

*Proof.* Set $C := \sqrt{-\log(|z|_p)/\log(R)}$ and let $n\geq 1$ be an integer. By (5-1), for a positive divisor $k$ of $n$ and $\ell\in\boldsymbol{D}(0,1)^*$ with $\ell^k = z^{n/k}$, we have

$$|\mathrm{Tate}(\ell)|_p = |\ell|_p^{-1} = |z|_p^{-n/k^2}.$$

Noting that $|z|_p^{-n/k^2} > R$ is equivalent to $k < Cn^{1/2}$, from (5-3) we deduce

$$\deg(T_n(\mathrm{Tate}(z))|_{\boldsymbol{D}^\infty(0,R)}) = \sum_{\substack{k>0,k\,|\,n \\ 0<k<C\sqrt{n}}} k < Cn^{1/2}d(n).$$

This completes the proof of the proposition. □

**5B.** ***Hecke orbits in the ordinary reduction locus.*** The goal of this section is to prove the following result describing, for an elliptic curve $E$ in $Y_{\mathrm{ord}}(\mathbb{C}_p)$, the asymptotic distribution of the Hecke orbit $(T_n(E))_{n=1}^\infty$. In the case where $E$ is in $Y_{\mathrm{ord}}(\mathbb{C}_p)$, Theorem C with $n=p^m n_0$ is a direct consequence of this result together with (2-1) and Lemma 2.3.

**Proposition 5.2.** *Let $\boldsymbol{D}$ be a residue disc contained in $Y_{\mathrm{ord}}(\mathbb{C}_p)$ and let $\boldsymbol{B}$ be a disc of radius strictly less than 1 contained in $Y_{\mathrm{ord}}(\mathbb{C}_p)$. Then for every $\varepsilon>0$ there is a constant $C>0$ such that for every $E$ in $\boldsymbol{D}$ and all integers $m\geq 0$ and $n_0\geq 1$ such that $n_0$ is not divisible by $p$, we have*

$$\deg(T_{p^m n_0}(E)|_{\boldsymbol{B}}) \leq C(m+1)n_0^{\varepsilon}.$$

To prove Proposition 5.2 we use the multiplicative property of the Hecke correspondences, see (2-6) in Section 2B. We first treat the case $n_0=1$ (Propositions 5.3) and the case $m=0$ (Propositions 5.4) separately. The proof of Proposition 5.2 is given at the end of this section.

**Proposition 5.3.** *Let $\zeta$ in $\bar{\mathbb{F}}_p$ be the $j$-invariant of an ordinary elliptic curve, denote by $r$ the minimal period of $\zeta$ under the Frobenius map $z\mapsto z^p$ and put $\boldsymbol{O} := \bigcup_{i=0}^{r-1}\boldsymbol{D}(\zeta^{p^i})$. Then for every $E$ in $\boldsymbol{D}(\zeta)$ and every integer $m\geq 1$, we have*

$$\mathrm{supp}(T_{p^m}(E)) \subseteq \boldsymbol{O}. \tag{5-4}$$

*Moreover, for every disc* $\boldsymbol{B}$ *of radius strictly less than* 1 *contained in* $\boldsymbol{O}$ *there is a constant* $C_1 > 0$ *such that for every* $E$ *in* $\boldsymbol{O}$ *and every integer* $m \geq 1$, *we have*

$$\deg(T_{p^m}(E)|_{\boldsymbol{B}}) \leq C_1 m. \tag{5-5}$$

*Proof.* The inclusion (5-4) is a direct consequence of Proposition 3.4 and (3-7). To prove (5-5), let $e$ be an ordinary elliptic curve with $j$-invariant $\zeta$, for every integer $i \geq 0$ put $z_i := \boldsymbol{t}^i(e^{\uparrow})$ and for every integer $i \leq -1$ let $i'$ be the unique integer in $\{0, \ldots, r-1\}$ such that $i - i'$ is divisible by $r$ and put $z_i := z_{i'}$. Note that for all nonnegative integers $a, b$, every integer $i$ and every point $z$ in $\boldsymbol{D}(z_i, 1)$, the set $\boldsymbol{t}^{-a}(\boldsymbol{t}^b(z))$ is contained in $\boldsymbol{D}(z_{i+b-a}, 1)$. Let $c$ in $]0, 1[$ be such that $\boldsymbol{B}$ is contained in $\boldsymbol{B}(c) := \bigcup_{i=0}^{r-1} \boldsymbol{D}(z_i, c)$, let $\rho$ and $\kappa_c$ be given by Lemma 3.7 and let $i_1 \geq 0$ be a sufficiently large integer so that $c\kappa_c^{i_1} < \rho$.

Fix $E$ in $\bigcup_{i=0}^{r-1} \boldsymbol{D}(z_i, 1)$ and let $m \geq 1$ be a given integer. Without loss of generality we assume $E \in \boldsymbol{D}(z_0, 1)$. We treat the cases $m < i_1$ and $m \geq i_1$ separately. If $m < i_1$, then we have

$$\deg(T_{p^m}(E)|_{\boldsymbol{B}(c)}) \leq \deg(T_{p^m}(E)) = \frac{p^{m+1} - 1}{p - 1} \leq p^{i_1} m.$$

Now, assume $m \geq i_1$. If for every $i$ in $\{0, \ldots, m\}$ the set $\boldsymbol{t}^{-(m-i)}(\boldsymbol{t}^i(E))$ is disjoint from $\boldsymbol{D}(z_{2i-m}, c)$, then

$$\deg(T_{p^m}(E)|_{\boldsymbol{B}(c)}) = \sum_{i=0}^{m} \deg((\boldsymbol{t}^*)^{(m-i)}([\boldsymbol{t}^i(E)])|_{\boldsymbol{D}(z_{2i-m}, c)}) = 0.$$

So we assume this is not the case and denote by $i_0$ the least integer $i$ in $\{0, \ldots, m\}$ such that $\boldsymbol{t}^{-(m-i)}(\boldsymbol{t}^i(E))$ contains a point $E_0$ in $\boldsymbol{D}(z_{2i-m}, c)$. Note that by Lemma 3.7(iii) the point $E_1 := \boldsymbol{t}^{i_1}(E_0)$ satisfies

$$|E_1 - z_{2i_0 - m + i_1}|_p \leq c\kappa_c^{i_1} < \rho,$$

so it is in $\boldsymbol{D}(z_{2i_0 - m + i_1}, \rho)$.

If $m \leq i_0 + i_1$, then we have

$$\deg(T_{p^m}(E)|_{\boldsymbol{B}(c)}) = \sum_{i=i_0}^{m} \deg((\boldsymbol{t}^*)^{m-i}([\boldsymbol{t}^i(E)])) \leq \sum_{i=i_0}^{m} p^{m-i} = \frac{p^{m-i_0+1} - 1}{p - 1} \leq p^{i_1}(m+1).$$

Suppose $m > i_0 + i_1$, and let $i$ be an integer satisfying $i_0 \leq i \leq m - i_1$. Noting that for every $E'$ in $\boldsymbol{t}^{-(m-i)}(\boldsymbol{t}^i(E))$ we have

$$\deg_{\boldsymbol{t}^{m-i}}(E') = \deg_{\boldsymbol{t}^{m-i-i_1}}(\boldsymbol{t}^{i_1}(E')) \deg_{\boldsymbol{t}^{i_1}}(E'),$$

we obtain

$$(\boldsymbol{t}^*)^{m-i}([\boldsymbol{t}^i(E)]) = \sum_{E'' \in \boldsymbol{t}^{-(m-i-i_1)}(\boldsymbol{t}^i(E))} \deg_{\boldsymbol{t}^{m-i-i_1}}(E'')(\boldsymbol{t}^*)^{i_1}([E'']). \tag{5-6}$$

On the other hand, for every $z$ in $\boldsymbol{t}^{-(m-i)}(\boldsymbol{t}^i(E))$ contained in $\boldsymbol{D}(z_{2i-m}, c)$, we have by Lemma 3.7(iii) and our choice of $i_1$,

$$|\boldsymbol{t}^{i_1}(z) - z_{2i-m+i_1}|_p \leq c\kappa_c^{i_1} < \rho,$$

so $\boldsymbol{t}^{i_1}(z) \in \boldsymbol{D}(z_{2i-m+i_1}, \rho)$. Since for such $z$ we have

$$\boldsymbol{t}^{m-i-i_1}(\boldsymbol{t}^{i_1}(z)) = \boldsymbol{t}^{i}(E) = \boldsymbol{t}^{m-i-i_1}(\boldsymbol{t}^{2i-2i_0}(E_1))$$

and by Lemma 3.7(ii) the map $\boldsymbol{t}^{m-i-i_1}$ is injective on $\boldsymbol{D}(z_{2i-m+i_1}, \rho)$, we conclude that $\boldsymbol{t}^{i_1}(z) = \boldsymbol{t}^{2i-2i_0}(E_1)$. Since we also have

$$\deg_{\boldsymbol{t}^{m-i-i_1}}(\boldsymbol{t}^{2i-2i_0}(E_1)) = 1$$

by Lemma 3.7(ii), when we restrict (5-6) to $\boldsymbol{D}(z_{2i-m}, c)$ we obtain

$$(\boldsymbol{t}^*)^{m-i}([\boldsymbol{t}^{i}(E)])|_{\boldsymbol{D}(z_{2i-m},c)} = (\boldsymbol{t}^*)^{i_1}([\boldsymbol{t}^{2i-2i_0}(E_1)])|_{\boldsymbol{D}(z_{2i-m},c)},$$

and therefore

$$\deg((\boldsymbol{t}^*)^{m-i}([\boldsymbol{t}^{i}(E)])|_{\boldsymbol{D}(z_{2i-m},c)}) \le \deg((\boldsymbol{t}^*)^{i_1}([\boldsymbol{t}^{2i-2i_0}(E_1)])) = p^{i_1}.$$

Together with Proposition 3.4 and our definition of $i_0$, this implies

$$\begin{aligned}\deg(T_{p^m}(E)|_{\boldsymbol{B}(c)}) &\le \sum_{i=i_0}^{m-i_1-1} \deg((\boldsymbol{t}^*)^{m-i}([\boldsymbol{t}^{i}(E)])|_{\boldsymbol{D}(z_{2i-m},c)}) + \sum_{i=m-i_1}^{m} \deg((\boldsymbol{t}^*)^{m-i}([\boldsymbol{t}^{i}(E)])) \\ &\le p^{i_1}(m-i_0-i_1) + \sum_{i=m-i_1}^{m} p^{m-i} \\ &\le p^{i_1}(m+1).\end{aligned}$$

This completes the proof of Proposition 5.3 with $C_1 = 2p^{i_1}$. □

**Proposition 5.4.** *Let $\boldsymbol{D}$ and $\boldsymbol{D}'$ be residue discs contained in $Y_{\mathrm{ord}}(\mathbb{C}_p)$. Then for every $\varepsilon > 0$ there is a constant $C_2 > 0$ such that for every $E$ in $\boldsymbol{D}$ and every integer $n \ge 1$ that is not divisible by $p$, we have*

$$\deg(T_n(E)|_{\boldsymbol{D}'}) \le C_2 n^{\varepsilon}.$$

To prove this proposition we first establish an intermediate estimate.

**Lemma 5.5.** *Let $e$ and $e'$ be ordinary elliptic curves over $\bar{\mathbb{F}}_p$, and for each integer $n \ge 1$ denote by $\mathrm{Hom}_n(e, e')$ the set of isogenies from $e$ to $e'$ of degree $n$. Then, for every $\varepsilon > 0$ we have*

$$\#\,\mathrm{Hom}_n(e, e') = o(n^{\varepsilon}). \tag{5-7}$$

*Proof.* Assume there is a nonzero element $\phi_0$ in $\mathrm{Hom}(e', e)$, for otherwise there is nothing to prove. Then, the map $\iota\colon\ \mathrm{Hom}(e, e') \to \mathrm{End}(e)$ given by $\iota(\phi) = \phi_0 \circ \phi$ is an injection, and $\deg(\iota(\phi)) = \deg(\phi_0)\deg(\phi)$. It is thus enough to prove (5-7) when $e' = e$.

Since $e$ is ordinary, the ring $\mathrm{End}(e)$ is isomorphic to an order inside a quadratic imaginary extension $K$ of $\mathbb{Q}$. Moreover, the isomorphism can be taken such that the degree of an isogeny is the same as the field norm of the corresponding element in $K$; see, e.g., [Silverman 2009, Chapter V, Theorem 3.1]. Let $d$ be the discriminant of $K$. Then $\mathcal{O}_{d,1}$ is the ring of integers of $K$, and hence it is enough to show

$$\#\{x \in \mathcal{O}_{d,1} : x\bar{x} = n\} = o(n^{\varepsilon}).$$

Since the group of units $\mathcal{O}_{d,1}^{\times}$ is finite, this estimate follows from (2-2) and (2-13). □

*Proof of Proposition 5.4.* Let $e$ be the ordinary elliptic curve over $\bar{\mathbb{F}}_p$ so that $\boldsymbol{D}' = \boldsymbol{D}(j(e))$. In view of Lemma 5.5, it is sufficient to show that for every $E$ in $\boldsymbol{D}$ and every integer $n \geq 1$ that is not divisible by $p$ we have

$$\deg(T_n(E)|_{\boldsymbol{D}'}) \leq \#\operatorname{Hom}_n(\tilde{E}, e). \tag{5-8}$$

Since the function $E \mapsto \deg(T_n(E)|_{\boldsymbol{D}'})$ is locally constant by Lemma 2.1, it is sufficient to establish this inequality in the case where $E$ is in $Y_{\mathrm{ord}}(\overline{\mathbb{C}_p^{\mathrm{unr}}})$.

To prove (5-8), recall that the reduction morphism $E \to \tilde{E}$ induces a bijective map $E[n] \to \tilde{E}[n]$; see for example [Silverman 2009, Chapter VII, Proposition 3.1(b)]. In addition, note that for a subgroup $C$ of $E$ of order $n$ such that $j(E/C)$ is in $\boldsymbol{D}'$, there is an isogeny $\tilde{E} \to e$ whose kernel is equal to the reduction of $C$. This defines an injective map

$$\{C \leq E : \#C = n,\ j(E/C) \in \boldsymbol{D}'\} \to \operatorname{Hom}_n(\tilde{E}, e),$$

proving (5-8) and completing the proof of the proposition. □

*Proof of Proposition 5.2.* Let $\zeta$ in $\bar{\mathbb{F}}_p$ be such that $\boldsymbol{B} \subseteq \boldsymbol{D}(\zeta)$, let $r \geq 1$ be the minimal period of $\zeta$ under the Frobenius map and put $\boldsymbol{O} := \bigcup_{i=0}^{r-1} \boldsymbol{D}(\zeta^{p^i})$. Let $C_1$ be given by Proposition 5.3 and let $C_2$ be the maximum value of the constants given by Proposition 5.4 with $\mathbf{D} = \mathbf{D}(\zeta), \ldots, \mathbf{D}(\zeta^{p^{r-1}})$.

Let $E$ in $\boldsymbol{D}$ be given. By (5-4), for every $E'$ in $\operatorname{supp}(T_{n_0}(E))$ that is not in $\boldsymbol{O}$ we have

$$\deg(T_{p^m}(E')|_{\boldsymbol{B}}) \leq \deg(T_{p^m}(E')|_{\boldsymbol{O}}) = 0.$$

On the other hand, for every $E'$ in $\operatorname{supp}(T_{n_0}(E))$ that is in $\boldsymbol{O}$, we have by Proposition 5.3

$$\deg(T_{p^m}(E')|_{\boldsymbol{B}}) \leq C_1 m + 1.$$

Together with (2-6) and Proposition 5.4 with $\boldsymbol{D}' = \boldsymbol{D}(\zeta), \ldots, \boldsymbol{D}(\zeta^{p^{r-1}})$, this implies

$$\deg(T_{p^m n_0}(E)|_{\boldsymbol{B}}) \leq (C_1 m + 1)\deg(T_{n_0}(E)|_{\boldsymbol{O}}) \leq r C_2 (C_1 + 1)(m+1) n_0^{\varepsilon}.$$

This proves the theorem with $C = r C_2 (C_1 + 1)$. □

**5C.** ***Hecke orbits in the supersingular reduction locus.*** The purpose of this section is to prove the following result on Hecke orbits inside the supersingular reduction locus. In the case where $E$ is in $Y_{\mathrm{sups}}(\mathbb{C}_p)$, Theorem C with $n = p^m n_0$ is a direct consequence of this result together with (2-1) and Lemma 2.3.

**Proposition 5.6.** *For every $e$ in $Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)$ fix an arbitrary $\gamma_e$ in $\boldsymbol{D}(j(e))$ and for every $r > 0$, put*

$$\boldsymbol{B}(r) := \bigcup_{e \in Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)} \boldsymbol{D}(\gamma_e, r).$$

*Then the following properties hold*:

(i) *For every $r$ in* $]0,1[$ *there is a constant $C>0$ such that for every $E$ in $Y_{\mathrm{sups}}(\mathbb{C}_p)$, every integer $m\geq 0$ and every integer $n_0\geq 1$ that is not divisible by $p$, we have*

$$\deg(T_{p^m n_0}(E)|_{\boldsymbol{B}(r)})\leq C\sigma_1(n_0).$$

(ii) *For every $r_0$ in* $]0,1[$ *and every integer $m_0\geq 0$, there is $r$ in* $]0,1[$ *such that for every $m$ in $\{0,\dots,m_0\}$ and integer $n_0\geq 1$ not divisible by $p$, we have for every $E$ in $\boldsymbol{B}(r_0)$*

$$\mathrm{supp}(T_{p^m n_0}(E))\subseteq \boldsymbol{B}(r).$$

The proof of this result is based on the following lemma, giving for each integer $m\geq 0$ a formula for the correspondence $\tau_m$ defined in Proposition 4.5. To state this lemma, for each integer $k\geq 0$ put

$$x_k:=\frac{p}{p+1}\cdot p^{-k}\quad\text{and}\quad I_k:=[x_{k+1},x_k],$$

and note that $\bigcup_{k=0}^{\infty} I_k=]0,\,p/(p+1)]$. Moreover, for all integers $k,k'\geq 0$ denote by

$$A^{(+1)}_{k,k'}\colon I_k\to I_{k'}\quad\text{and}\quad A^{(-1)}_{k,k'}\colon I_k\to I_{k'}$$

the unique affine bijection preserving or reversing the orientation, respectively. Note that for every $k\geq 0$ we have $1-A^{(+1)}_{k,0}=A^{(-1)}_{k,0}$ and that for every $k'\geq 1$ we have

$$pA^{(\pm 1)}_{k,k'}=A^{(\pm 1)}_{k,k'-1}. \tag{5-9}$$

**Lemma 5.7.** *For each integer $m\geq 0$ denote by $\tau_m$ the correspondence acting on $[0,\,p/(p+1)]$ defined in Proposition 4.5. Then for all integers $k,m\geq 0$, we have*

$$\tau_m|_{I_k}=\begin{cases}\sum_{i=0}^{m} p^i (A^{(+1)}_{k,2i-(m-k)})_* & \text{if } m\leq k;\\ \sum_{i=0}^{m-k-1} p^i (A^{((-1)^{m-k-i})}_{k,i})_*+\sum_{i=m-k}^{m} p^i (A^{(+1)}_{k,2i-(m-k)})_* & \text{if } m\geq k+1.\end{cases}$$

*Proof.* Fix $k\geq 0$. We proceed by induction on $m$. The case $m=0$ is trivial and the case $m=1$ is a direct consequence of the definition given in Proposition 4.5. Let $m\geq 2$ be given and suppose that the lemma holds with $m$ replaced by $m-1$ and by $m-2$. If $m\leq k$, then by (5-9)

$$\tau_1(\tau_{m-1}|_{I_k})=\sum_{i=0}^{m-1} p^i (A^{(+1)}_{k,2i-(m-k)})_*+\sum_{i=0}^{m-1} p^{i+1}(A^{(+1)}_{k,2i-(m-k)+2})_*=p\tau_{m-2}|_{I_k}+\sum_{i=0}^{m} p^i (A^{(+1)}_{k,2i-(m-k)})_*,$$

which proves the induction step in the case $m\leq k$. In the case $m=k+1$, using $1-A^{(+1)}_{k,0}=A^{(-1)}_{k,0}$ we have

$$\tau_1(\tau_k|_{I_k})=(A^{(-1)}_{k,0})_*+\sum_{i=1}^{k} p^i (A^{(+1)}_{k,2i-1})_*+\sum_{i=0}^{k} p^{i+1}(A^{(+1)}_{k,2i+1})_*=p\tau_{k-1}|_{I_k}+(A^{(-1)}_{k,0})_*+\sum_{i=1}^{k+1} p^i (A^{(+1)}_{k,2i-1})_*.$$

This proves the induction step in the case $m = k+1$. If $m = k+2$, then

$$\begin{aligned}\tau_1(\tau_{k+1}|_{I_k}) &= (A_{k,0}^{(+1)})_* + p(A_{k,1}^{(-1)})_* + \sum_{i=1}^{k+1} p^i (A_{k,2i-2}^{(+1)})_* + \sum_{i=1}^{k+1} p^{i+1} (A_{k,2i}^{(+1)})_* \\ &= (A_{k,0}^{(+1)})_* + p(A_{k,1}^{(-1)})_* + p\tau_k|_{I_k} + \sum_{i=2}^{k+2} p^i (A_{k,2i-2}^{(+1)})_*.\end{aligned}$$

This proves the induction step in the case $m = k+2$. Finally, if $m \geq k+3$, then $\tau_1(\tau_{m-1}|_{I_k})$ is equal to

$$\begin{aligned}&(A_{k,0}^{((-1)^{m-k})})_* + \sum_{j=1}^{m-k-2} p^j (A_{k,j-1}^{((-1)^{m-k-j-1})})_* + \sum_{j=0}^{m-k-2} p^{j+1} (A_{k,j+1}^{((-1)^{m-k-j-1})})_* + \sum_{i=m-k-1}^{m-1} p^i (A_{k,2i-(m-k)}^{(+1)})_* \\ &\qquad + \sum_{i=m-k-1}^{m-1} p^{i+1} (A_{k,2i-(m-k-2)}^{(+1)})_* \\ &= \sum_{\ell=0}^{m-k-1} p^\ell (A_{k,\ell}^{((-1)^{m-k-\ell})})_* + p \sum_{s=0}^{m-k-3} p^s (A_{k,s}^{((-1)^{m-k-s-2})})_* + \sum_{i=m-k}^{m} p^i (A_{k,2i-(m-k)}^{(+1)})_* \\ &\qquad + p \sum_{i=m-k-2}^{m-2} p^i (A_{k,2i-(m-k-2)}^{(+1)})_* \\ &= p\tau_{m-2}|_{I_k} + \sum_{\ell=0}^{m-k-1} p^\ell (A_{k,\ell}^{((-1)^{m-k-\ell})})_* + \sum_{i=m-k}^{m} p^i (A_{k,2i-(m-k)}^{(+1)})_*.\end{aligned}$$

This completes the proof of the induction step and of the lemma. □

*Proof of Proposition 5.6.* Let $\hat{v}_p$ and $(\tau_m)_{m=0}^{\infty}$ be as in Proposition 4.5.

To prove (i), let $r$ in $]0, 1[$ be given. By Proposition 4.3 there is an integer $\ell \geq 0$ such that $\hat{v}_p(\boldsymbol{B}(r)) \subseteq [x_\ell, x_0]$. Then the desired assertion follows from Proposition 4.5 and by the observation that by Lemma 5.7 for every $x$ in $]0, x_0]$ we have

$$\deg(\tau_m(x)|_{[x_\ell, x_0]}) \leq 1 + p + \cdots + p^\ell.$$

To prove (ii), let $r_0$ in $]0, 1[$ and an integer $m_0 \geq 0$ be given. By Proposition 4.3 there is an integer $\ell \geq 0$ such that $\hat{v}_p(\boldsymbol{B}(r_0)) \subseteq [x_\ell, x_0]$ and $r$ in $]0, 1[$ such that $\hat{v}_p^{-1}([x_{\ell+m_0}, x_0]) \subseteq \boldsymbol{B}(r)$. Then the desired inclusion follows from Proposition 4.5 by noting that by Lemma 5.7 for every $x$ in $[x_\ell, x_0]$ and every $m$ in $\{0, \ldots, m_0\}$, we have $\operatorname{supp}(\tau_m(x)) \subseteq [x_{\ell+m_0}, x_0]$. □

## Appendix A: Lifting the Hasse invariant in characteristic 2 and 3

When $p$ equals 2 or 3 it is not possible to lift the Hasse invariant $A_{p-1}$ to a modular form of level one, holomorphic at infinity, over $\mathbb{Z}_{(p)}$. There are two approaches to solve this issue. On the one hand, there are liftings of $A_1^4$ and $A_2^3$ in the desired space (namely, the Eisenstein series $\mathrm{E}_4$ and $\mathrm{E}_6$). On the other hand, considering level structures, liftings can be constructed as algebraic modular forms over $\mathbb{Z}_{(p)}$ of the expected weight but higher level. In this appendix we recall both approaches, following [Katz 1973,

Section 2.1], and give a quantitative comparison between them, embodied in Proposition A.1 below. Such comparison is needed in Section 4B.

We start by recalling level structures. Let $R$ be a ring and let $n \geq 1$ be an integer which is assumed to be invertible in $R$. Let $E$ be an elliptic curve over $R$ in the sense of Section 4A. A *level n structure* on $E$ over $R$ is an isomorphism $\alpha_n \colon E[n] \to (\mathbb{Z}/n\mathbb{Z})^2$ of group schemes over $R$.

Given an integer $n \geq 1$ and an arbitrary ring $R_0$ where $n$ is invertible, an algebraic modular form of level $n \geq 1$ over $R_0$ is a family of maps $F = (F_R)_{R \in R_0\text{-Alg}}$ such that for any $R \in R_0$-Alg, the $R$-valued map $F_R$ is defined on the set of triples $(E, \omega, \alpha_n)$, where $E$ is an elliptic curve over $R \in R_0$-Alg, together with a differential form in $\Omega^1_{E/R}(E)'$ and a level $n$ structure. The element $F_R(E, \omega, \alpha_n) \in R$ must define an assignment satisfying properties analogous to (i), (ii) and (iii) stated in Section 4A. See [Katz 1973, Section 1.2] for further details.

When $R_0$ contains $1/n$ and a primitive $n$-th root of unity, the $q$-expansions of an algebraic modular form $F$ of level $n$ over $R_0$ are defined as the elements of $\mathbb{Z}((q)) \otimes_{\mathbb{Z}} R_0$ obtained by evaluating $F$ at the triples $(\mathrm{Tate}(q^n), \omega_{\mathrm{can}}, \alpha_n)_{R_0}$ consisting of the Tate curve $\mathrm{Tate}(q^n)$ (see Section 5A) with its canonical differential $\omega_{\mathrm{can}}$, regarded over $\mathbb{Z}((q)) \otimes_{\mathbb{Z}} R_0$, with $\alpha_n$ varying over all level $n$ structures of $\mathrm{Tate}(q^n)$ over $\mathbb{Z}((q)) \otimes_{\mathbb{Z}} R_0$. If all of the $q$-expansions of $F$ lie in $\mathbb{Z}[\![q]\!] \otimes_{\mathbb{Z}} R_0$ then $F$ is called holomorphic at infinity. For algebraic modular forms $F$ of level one there is only one $q$-expansion, which coincides with the previously defined $F(q)$.

According to [Katz 1973, page 98], for any level $3 \leq n \leq 11$ odd, there exists a lifting of $A_1$ to a modular form of level $n$, weight one, holomorphic at infinity, over $\mathbb{Z}[1/n]$. We define $\mathrm{E}_1$ as any such lifting and set $n(\mathrm{E}_1) := n$. Similarly, when $m \geq 4$ and $3 \nmid m$, there exists a lifting of $A_2$ to a modular form of level $m$, weight two, holomorphic at infinity, over $\mathbb{Z}[1/m]$. We define $\mathrm{E}_2$ as any such lifting and set $n(\mathrm{E}_2) := m$.

The following statement is a comparison between both approaches.

**Proposition A.1.** *Let $E \in Y_{\mathrm{sups}}(\mathbb{C}_p)$ and let $\omega$ be a differential form in $\Omega^1_{E/\mathcal{O}_p}(E)'$:*

(i) *For any level $n(\mathrm{E}_1)$ structure $\alpha$ on $E$ we have*

$$\mathrm{ord}_2(\mathrm{E}_4(E, \omega)) < 3 \Leftrightarrow \mathrm{ord}_2(\mathrm{E}_1^4(E, \omega, \alpha)) < 3,$$

*in which case* $\mathrm{ord}_2(\mathrm{E}_4(E, \omega)) = \mathrm{ord}_2(\mathrm{E}_1^4(E, \omega, \alpha))$.

(ii) *For any level $n(\mathrm{E}_2)$ structure $\alpha$ on $E$ we have*

$$\mathrm{ord}_3(\mathrm{E}_6(E, \omega)) < \tfrac{5}{2} \Leftrightarrow \mathrm{ord}_3(\mathrm{E}_2^3(E, \omega, \alpha)) < \tfrac{5}{2},$$

*in which case* $\mathrm{ord}_3(\mathrm{E}_4(E, \omega)) = \mathrm{ord}_3(\mathrm{E}_2^3(E, \omega, \alpha))$.

*Proof.* In order to prove (i), we start by recalling the $q$-expansion

$$\mathrm{E}_4(q) = 1 + 240 \sum_{n=1}^{\infty} \sigma_3(n) q^n,$$

obtained by setting $k=4$ in (4-2). Since $\mathrm{ord}_2(240)=4$, we have $\mathrm{E}_4(q)\equiv 1 \bmod 2^4$. Now, put $n_1 := n(\mathrm{E}_1)$, let $\zeta_{n_1}$ be a primitive $n_1$-th roof of unity and define $R_1 := \mathbb{Z}[1/n_1, \zeta_{n_1}]$. By the definition of $\mathrm{E}_1$ we have

$$\mathrm{E}_1(\mathrm{Tate}(q^{n_1}), \omega_{\mathrm{can}}, \alpha_{n_1}) \equiv A_1(q^{n_1}) \equiv 1 \bmod 2R_1,$$

hence

$$\mathrm{E}_1^4(\mathrm{Tate}(q^{n_1}), \omega_{\mathrm{can}}, \alpha_{n_1}) \equiv 1 \equiv \mathrm{E}_4(q^{n_1}) \bmod 2^3 R_1,$$

for any level $n_1$ structure $\alpha_{n_1}$ on $\mathrm{Tate}(q^{n_1})$. We conclude that the form $f$ obtained by reducing modulo $2^3\mathbb{Z}[1/n_1]$ the form $\mathrm{E}_4 - \mathrm{E}_1^4$ is an algebraic modular form of weight 4, level $n_1$ over $\mathbb{Z}/2^3\mathbb{Z}$, whose $q$-expansions over $(\mathbb{Z}/2^3\mathbb{Z})[\zeta_{n_1}]$ vanish identically. By [Katz 1973, Theorem 1.6.1] we deduce that $f=0$. By compatibility with base change we conclude that for any $\mathbb{Z}[1/n_1]$-algebra $R$ and any triple $(E,\omega,\alpha_{n_1})$ over $R$ we have

$$\mathrm{E}_4(E,\omega) - \mathrm{E}_1^4(E,\omega,\alpha_{n_1}) \equiv f((E,\omega,\alpha_{n_1})_{R/2^3R}) \equiv 0 \bmod 2^3 R.$$

In particular, choosing $R=\mathcal{O}_p$, we get

$$\mathrm{ord}_2(\mathrm{E}_4(E,\omega) - \mathrm{E}_1^4(E,\omega,\alpha_{n_1})) \geq 3, \tag{A-1}$$

for every $E \in Y_{\mathrm{sups}}(\mathbb{C}_p)$, every basis $\omega$ of $\Omega^1_{E/\mathcal{O}_p}$ and every level $n_1$ structure $\alpha_{n_1}$ on $E$. Then, (i) is a direct consequence of (A-1) and the ultrametric inequality.

The proof of (ii) is unfortunately less straightforward. This is because the same argument used to prove (A-1) only yields the inequality

$$\mathrm{ord}_3(\mathrm{E}_6(E,\omega) - \mathrm{E}_2^3(E,\omega,\alpha_{n_2})) \geq 2,$$

valid for any level $n_2 := n(\mathrm{E}_2)$ structure $\alpha_{n_2}$ on $E$, but such inequality does not imply the desired result. On the other hand, the above argument allows us to infer

$$\mathrm{ord}_3(\mathrm{E}_4(E,\omega) - \mathrm{E}_2^2(E,\omega,\alpha_{n_2})) \geq 1. \tag{A-2}$$

In order to prove (ii) we introduce the series

$$G_2(\tau) = 1 + 24\sum_{n=1}^{\infty}\left(\sigma_1(n) - 2\sigma_1\left(\frac{n}{2}\right)\right)e^{2\pi i n\tau}, \quad \tau\in\mathbb{H}, \tag{A-3}$$

where $\sigma_1\left(\frac{n}{2}\right)$ is defined as zero when $n$ is odd. It is known that $G_2$ is a classical holomorphic modular form of weight two for the group $\Gamma_0(2) = \left\{g\in \mathrm{SL}_2(\mathbb{Z}) : g \equiv \left(\begin{smallmatrix} * & * \\ 0 & * \end{smallmatrix}\right) \bmod 2\right\}$.[‖] By [Katz 1973, Corollary 1.9.1], $G_2$ defines an algebraic modular over $\mathbb{Z}\left[\frac{1}{2}\right]$ of weight two and level two. This form satisfies the identity

$$4G_2^3 = \mathrm{E}_6 + 3\,\mathrm{E}_4\,G_2. \tag{A-4}$$

Indeed, the space of modular forms over $\mathbb{C}$ of weight six for $\Gamma_0(2)$ has dimension 2, see the dimension formulas in [Diamond and Shurman 2005, Chapter 3]. By comparing Fourier expansions, it is easy

[‖] Up to an explicit multiplicative factor, this is denoted by $G_{2,2}$ in [Diamond and Shurman 2005, Section 1.2].

to check that $\mathrm{E}_6$ and $\mathrm{E}_4 G_2$ are linearly independent over $\mathbb{C}$, hence they form a basis of such space. This implies that there exist $a, b \in \mathbb{C}$ with $G_2^3 = a\,\mathrm{E}_6 + b\,\mathrm{E}_4 G_2$. Then, (A-4) follows at the level of classical modular forms by computing the values of $a$ and $b$, which can be done by comparing Fourier expansions. Finally, the fact that (A-4) holds as an identity between algebraic modular forms over $\mathbb{Z}\left[\frac{1}{2}\right]$ is a consequence of [Katz 1973, Corollary 1.9.1].

We also recall the identity

$$\mathrm{E}_6^2 - \mathrm{E}_4^3 = 1728\Delta.$$

At the level of classical modular forms, see for example [Diamond and Shurman 2005, Sections 1.1 and 1.2]. Then, this identity holds at the level of algebraic modular forms by the same reasoning as before. Given $E \in Y_{\mathrm{sups}}(\mathbb{C}_p)$ and a differential form $\omega$ in $\Omega^1_{E/\mathcal{O}_p}(E)'$, we have $\Delta(E, \omega) \in \mathcal{O}_p^\times$ since $E$ has good reduction. This implies

$$\mathrm{ord}_3(\mathrm{E}_6^2(E, \omega) - \mathrm{E}_4^3(E, \omega)) = 3. \tag{A-5}$$

By using (A-4) and (A-5), we will now prove (ii). Let $\alpha$ be a level $n_2$ structure on $E$. First, assume that $\mathrm{ord}_3(\mathrm{E}_2(E, \omega, \alpha)) < \frac{5}{6}$. From (A-4) we see that the reduction modulo 3 of $G_2$ equals $A_2$. Since the same holds for $\mathrm{E}_2$, we conclude that

$$\mathrm{ord}_3(\mathrm{E}_2(E, \omega, \alpha) - G_2(E, \omega, \beta)) \geq 1, \tag{A-6}$$

for any level two structure $\beta$. In particular

$$\mathrm{ord}_3(G_2(E, \omega, \beta)) = \mathrm{ord}_3(\mathrm{E}_2(E, \omega, \alpha)) < \tfrac{5}{6}.$$

By (A-4) we have

$$\mathrm{E}_6(E, \omega) = G_2(E, \omega, \beta)(4G_2^2(E, \omega, \beta) - 3\,\mathrm{E}_4(E, \omega)).$$

But by (A-2) and (A-6) we also have

$$\mathrm{ord}_3(3\,\mathrm{E}_4(E, \omega)) = 1 + \mathrm{ord}_3(\mathrm{E}_4(E, \omega)) \geq 1 + \min\{1, \mathrm{ord}_3(G_2^2(E, \omega, \beta))\} > \mathrm{ord}_3(G_2^2(E, \omega, \beta)),$$

hence

$$\mathrm{ord}_3(\mathrm{E}_6(E, \omega)) = \mathrm{ord}_3(G_2^3(E, \omega, \beta)) = \mathrm{ord}_3(\mathrm{E}_2^3(E, \omega, \alpha)).$$

This proves one implication. Let us now prove the reciprocal. We start by assuming that $\mathrm{ord}_3(\mathrm{E}_6(E, \omega)) < \frac{5}{2}$. If $\mathrm{ord}_3(\mathrm{E}_4(E, \omega)) < 1$, then we can use (A-2), (A-5) and (A-6) to deduce that $\mathrm{ord}_3(\mathrm{E}_4^3(E, \omega)) = \mathrm{ord}_3(\mathrm{E}_6^2(E, \omega))$ and $\mathrm{ord}_3(G_2^2(E, \omega, \beta)) = \mathrm{ord}_3(\mathrm{E}_4(E, \omega))$. This implies

$$\mathrm{ord}_3(3G_2(E, \omega, \beta)\,\mathrm{E}_4(E, \omega)) = 1 + \mathrm{ord}_3(\mathrm{E}_6(E, \omega)) > \mathrm{ord}_3(\mathrm{E}_6(E, \omega)).$$

By (A-4) and (A-6) we conclude

$$\mathrm{ord}_3(\mathrm{E}_2^3(E, \omega, \alpha)) = \mathrm{ord}_3(G_2^3(E, \omega, \beta)) = \mathrm{ord}_3(\mathrm{E}_6(E, \omega)).$$

Now, if $\operatorname{ord}_3(\mathrm{E}_4(E,\omega)) \geq 1$ then (A-2) and (A-6) imply $\operatorname{ord}_3(G_2^2(E,\omega,\beta)) \geq 1$, giving

$$\operatorname{ord}_3(3G_2(E,\omega,\beta)\,\mathrm{E}_4(E,\omega)) \geq \tfrac{5}{2} > \operatorname{ord}_3(\mathrm{E}_6(E,\omega)).$$

As before, we conclude $\operatorname{ord}_3(\mathrm{E}_2^3(E,\omega,\alpha)) = \operatorname{ord}_3(\mathrm{E}_6(E,\omega))$. This proves the reciprocal implication and completes the proof of the proposition. □

## Appendix B: Eichler–Shimura analytic relation

In this appendix we further study the canonical branch $\boldsymbol{t}$ of $T_p$ that is defined on $Y_{\mathrm{ord}}(\mathbb{C}_p)$ in Section 3A. We start extending $\boldsymbol{t}$, as follows. Recall that $v_p$ denotes Katz' valuation, defined in Section 4A. Extend $v_p$ to $Y(\mathbb{C}_p)$ as $v_p \equiv 0$ outside $Y_{\mathrm{sups}}(\mathbb{C}_p)$, and put

$$N_p := \left\{ E \in Y(\mathbb{C}_p) : v_p(E) < \frac{p}{p+1} \right\}. \tag{B-1}$$

On $N_p \cap Y_{\mathrm{sups}}(\mathbb{C}_p)$, we use the definition of $\boldsymbol{t}$ in Lemma 4.6. To define $\boldsymbol{t}$ at a point $E$ in $Y_{\mathrm{bad}}(\mathbb{C}_p)$, let $z$ in $\boldsymbol{D}(0,1)^*$ and let $\varphi_z\colon \mathbb{C}_p^\times/z^{\mathbb{Z}} \to \mathrm{Tate}(z)(\mathbb{C}_p)$ be the isomorphism of analytic groups as in Section 5A. Then we define

$$H(E) := \varphi_z(\{\zeta z^n \in \mathbb{C}_p^\times : \zeta^p = 1, n \in \mathbb{Z}\}/z^{\mathbb{Z}}), \quad \text{and} \quad \boldsymbol{t}(E) := E/H(E).$$

Note that in the notation (5-2) of Section 5A, we have $H(E) = C_{p,z^p}$. The map $\boldsymbol{t}\colon N_p \to Y(\mathbb{C}_p)$ so defined is the *canonical branch of* $T_p$.

The goal of this appendix is to prove the following result.

**Theorem B.1** (Eichler–Shimura analytic relation). *The canonical branch $\boldsymbol{t}$ of $T_p$ is given by a finite sum of Laurent series, each of which converges on all of $N_p$. Furthermore, for every $E$ in $N_p \setminus Y_{\mathrm{bad}}(\mathbb{C}_p)$ we have*

$$\operatorname{ord}_p(\boldsymbol{t}(j(E)) - j(E)^p) \geq 1 - v_p(E), \tag{B-2}$$

*and for every $E$ in $Y(\mathbb{C}_p)$ we have*

$$T_p(E) = \begin{cases} \boldsymbol{t}^*(E) + [\boldsymbol{t}(E)] & \text{if } v_p(E) \leq 1/(p+1); \\ \boldsymbol{t}^*(E) & \text{if } v_p(E) > 1/(p+1). \end{cases} \tag{B-3}$$

In view of (B-2), the relation (B-3) can be seen as refinement and a lift to $N_p$ of the classical Eichler–Shimura congruence relation; see for example [Shimura 1971, Section 7.4] or [Diamond and Shurman 2005, Section 8.7].

The proof of Theorem B.1 is at the end of this appendix. When restricted to $Y_{\mathrm{ord}}(\mathbb{C}_p)$, it is a direct consequence of Theorem 3.3 and Proposition 3.4 with $m = 1$. To prove (B-3) for $E$ in $Y_{\mathrm{sups}}(\mathbb{C}_p)$, we use Lemma 4.6. To prove this relation on $Y_{\mathrm{bad}}(\mathbb{C}_p)$, we use the results on the uniformization of $p$-adic elliptic curves with multiplicative reduction, recalled in Section 5A. To prove (B-2) and that $\boldsymbol{t}$ is a finite sum of Laurent series for $p \geq 5$, we use Theorem 3.3 in Section 3A. For $p = 2$ and 3, we use Proposition B.2

below, whose proof is based on the explicit formulae in [Mestre 1986, Appendice]. This result also provides a proof of Theorem 3.3 when $p = 2$ and 3

Note that for $p = 2$ and 3, the set $Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)$ consists of a single point whose $j$-invariant is equal to 0 and to 1728; see for example [Silverman 2009, Chapter V, Section 4].

**Proposition B.2.** *Put* $\mathrm{j}_2 := 0$ *and* $\mathrm{j}_3 := 1728$, *and consider the polynomials*

$$\check{k}_2(z) := -93 \cdot 2^4 z + 627 \cdot 2^8 \quad \textit{and} \quad \check{k}_3(z) := 328 \cdot 3^2 z^2 + 85708 \cdot 3^3 z + 1263704 \cdot 3^5.$$

*Then for* $p = 2$ *and* 3, *the canonical branch* $\boldsymbol{t}$ *of* $T_p$ *admits a Laurent series expansion of the form*

$$\boldsymbol{t}(z) = (z - \mathrm{j}_p)^p + \mathrm{j}_p + \check{k}_p(z - \mathrm{j}_p) + \sum_{n=1}^{\infty} \frac{A_n^{(p)}}{(z - \mathrm{j}_p)^n},$$

*where for every* $n \geq 1$ *the coefficient* $A_n^{(p)}$ *is in* $\mathbb{Z}$ *and satisfies*

$$\mathrm{ord}_p(A_n^{(p)}) \geq \begin{cases} 4 + 8n & \text{if } p = 2; \\ \frac{3}{2} + \frac{9}{2}n & \text{if } p = 3, \end{cases}$$

*with equality if* $n = 1$.

To prove this proposition, we introduce some notation and recall the explicit formulae in [Mestre 1986, Appendice]. For $\mathbb{K} = \mathbb{C}$ or $\mathbb{C}_p$, we use $j$ to identify $Y(\mathbb{K})$ with $\mathbb{K}$ and consider $T_p$ as a correspondence acting on $\mathrm{Div}(\mathbb{K})$. Let $Y_0(p)$, $\alpha_p$ and $\beta_p$ be as in Section 2B, so that $T_p = (j \circ \alpha_p)_* \circ (j \circ \beta_p)^*$. Denote by

$$w_p \colon Y_0(p)(\mathbb{K}) \to Y_0(p)(\mathbb{K})$$

the *Atkin–Lehner* or *Fricke involution*, defined by $w_p(E, C) := (E/C, E[p]/C)$ and note that $\beta_p = \alpha_p \circ w_p$. Identify $Y_0(p)(\mathbb{C})$ with the quotient $\Gamma_0(p)\backslash\mathbb{H}$ and denote by $\eta \colon \mathbb{H} \to \mathbb{C}$ *Dedekind's eta function*, defined by

$$\eta(\tau) := \exp\left(\frac{\pi i \tau}{12}\right) \prod_{n=1}^{\infty} (1 - \exp(2\pi i n \tau)).$$

Then for $p = 2$ or 3, the function $\hat{x}_p \colon \mathbb{H} \to \mathbb{C}$ defined by

$$\hat{x}_p(\tau) := \left(\frac{\eta(\tau)}{\eta(p\tau)}\right)^{24/(p-1)}$$

descends to a complex analytic isomorphism $x_p \colon Y_0(p)(\mathbb{C}) \to \mathbb{C}$. Moreover, defining

$$\hat{\alpha}_p(z) := \begin{cases} (z + 2^4)^3/z & \text{if } p = 2; \\ (z + 3^3)(z + 3)^3/z & \text{if } p = 3, \end{cases} \quad \text{and} \quad \hat{w}_p(z) := \begin{cases} 2^{12}/z & \text{if } p = 2; \\ 3^6/z & \text{if } p = 3, \end{cases}$$

we have $j \circ \alpha_p = \hat{\alpha}_p \circ x_p$ and $x_p \circ w_p = \hat{w}_p \circ x_p$; see [Mestre 1986, pages 238 and 239]. It follows that, if we put

$$\hat{\beta}_p(z) := \hat{\alpha}_p \circ \hat{w}_p(z) = \begin{cases} (z + 2^8)^3/z^2 & \text{if } p = 2; \\ (z + 3^3)(z + 3^5)^3/z^3 & \text{if } p = 3, \end{cases}$$

then $j \circ \beta_p = \hat{\beta}_p \circ x_p$ and therefore $T_p = (\hat{\alpha}_p)_* \circ \hat{\beta}_p^*$ as algebraic correspondences over $\mathbb{C}$. Since $T_p$, $\hat{\alpha}_p$ and $\hat{\beta}_p$ are all defined over $\mathbb{Q}$, we have that the equality $T_p = (\hat{\alpha}_p)_* \circ \hat{\beta}_p^*$ also holds as algebraic correspondences over $\mathrm{Div}(Y(\mathbb{C}_p))$.

The following elementary lemma is used the proof of Proposition B.2. Given $r$ in $]0, 1[$, and a Laurent series $\sum_{n=0}^{\infty} \frac{A_n}{z^n}$ in $\mathbb{Z}\left[\!\left[\frac{1}{z}\right]\!\right]$, put

$$\left\| \sum_{n=0}^{\infty} \frac{A_n}{z^n} \right\|_r := \sup\{|A_n|_p r^{-n} : n \geq 0\}.$$

**Lemma B.3.** *Let $\delta(z)$ in $\frac{1}{z}\mathbb{Z}\left[\!\left[\frac{1}{z}\right]\!\right]$ be given and put $f(z) := z(1 + \delta(z))$. Then there is $\Delta(z)$ in $\frac{1}{z}\mathbb{Z}\left[\!\left[\frac{1}{z}\right]\!\right]$ such that $F(z) := z(1 + \Delta(z))$ satisfies $F(f(z)) = z$. If in addition for some $r$ in $]0, 1[$ we have $\|\delta\|_r \leq 1$, then $\|\Delta\|_r \leq 1$.*

*Proof.* We start defining recursively a sequence $(\Delta_n)_{n=0}^{\infty}$ in $\frac{1}{z}\mathbb{Z}\left[\frac{1}{z}\right]$ such that for every integer $n \geq 0$,

$$z^n \Delta_n(z) \in \mathbb{Z}[z], \quad \Delta_{n+1}(z) \equiv \Delta_n(z) \bmod \tfrac{1}{z^{n+1}}\mathbb{Z}\left[\tfrac{1}{z}\right],$$

and the Laurent polynomial $F_n(z) := z(1 + \Delta_n(z))$ satisfies

$$F_n(f(z)) \equiv z \bmod \tfrac{1}{z^n}\mathbb{Z}\left[\tfrac{1}{z}\right].$$

For $n = 0$ put $\Delta_0(z) = 0$, so $F_0(f(z)) = f(z) \equiv z \bmod \mathbb{Z}\left[\frac{1}{z}\right]$. Let $n \geq 0$ be an integer so that $\Delta_n$ is already defined and let $A$ in $\mathbb{Z}$ be the coefficient of $1/z^n$ in $F_n(f(z))$. Then for $\Delta_{n+1}(z) := \Delta_n(z) - A/z^{n+1}$, we have

$$(F_{n+1} - F_n)(f(z)) = -\frac{A}{z^n(1+\delta(z))^n} = -\frac{A}{z^n}\left(1 + \sum_{k=1}^{\infty}(-\delta(z))^k\right)^n \equiv -\frac{A}{z^n} \bmod \frac{1}{z^{n+1}}\mathbb{Z}\left[\frac{1}{z}\right],$$

and therefore

$$F_{n+1}(f(z)) - z = F_n(f(z)) - z + (F_{n+1} - F_n)(f(z)) \equiv 0 \bmod \tfrac{1}{z^{n+1}}\mathbb{Z}\left[\tfrac{1}{z}\right].$$

This completes the definition of the sequence $(\Delta_n)_{n=0}^{\infty}$. It follows that the unique series $\Delta$ in $\frac{1}{z}\mathbb{Z}\left[\!\left[\frac{1}{z}\right]\!\right]$ satisfying for every $n \geq 0$ the congruence

$$\Delta(z) \equiv \Delta_n(z) \bmod \tfrac{1}{z^{n+1}}\mathbb{Z}\left[\!\left[\tfrac{1}{z}\right]\!\right],$$

satisfies $F(f(z)) = z$.

To prove the last assertion, note that for every $r$ in $]0, 1[$,

$$I_r := \left\{z(1 + g(z)) : g(z) \in \tfrac{1}{z}\mathbb{Z}\left[\!\left[\tfrac{1}{z}\right]\!\right], \|g\|_r \leq 1\right\}$$

is a collection of series in $\mathbb{Z}\left[\!\left[\frac{1}{z}\right]\!\right]$ that is closed under composition. It follows from the above construction that, if for some $r$ in $]0, 1[$ we have $\|\delta\|_r \leq 1$, then for every integer $n \geq 0$ the series $F_n$ and $F_n \circ f$ are both in $I_r$. This implies that $F$ is in $I_r$, as wanted. □

The proof of Proposition B.2 is given after the following lemma, which is also used in the proof of Theorem B.1.

**Lemma B.4.** *For an arbitrary prime number $p$, the right-hand side of* (3-3) *converges to $\boldsymbol{t}$ on $Y_{\mathrm{ord}}(\mathbb{C}_p) \cup Y_{\mathrm{bad}}(\mathbb{C}_p)$.*

*Proof.* Let $\Phi_p(X, Y)$ be the modular polynomial of level $p$, as defined in Section 2B, so that for every $z$ in $Y_{\mathrm{ord}}(\mathbb{C}_p)$ we have $\Phi_p(z, \boldsymbol{t}(z)) = 0$. By Theorem 3.3, the finite sum of Laurent series on the right-hand side of (3-3) converges on $Y_{\mathrm{ord}}(\mathbb{C}_p) \cup Y_{\mathrm{bad}}(\mathbb{C}_p)$ to a function $\hat{\boldsymbol{t}}$ extending $\boldsymbol{t}$, and for $z$ in $Y_{\mathrm{bad}}(\mathbb{C}_p)$ we have $|\hat{\boldsymbol{t}}(z)|_p = |z|_p^p$. It follows that for every $z$ in $Y_{\mathrm{bad}}(\mathbb{C}_p)$ we have $\Phi_p(z, \hat{\boldsymbol{t}}(z)) = 0$, so $\hat{\boldsymbol{t}}(z)$ is in the support of $T_p(z)$. Combining (5-1) and (5-3), we conclude that $\hat{\boldsymbol{t}}(z) = \boldsymbol{t}(z)$. □

*Proof of Proposition B.2.* Note that if we put $r_2 := 2^{-8}$ and $r_3 := 3^{-9/2}$, then for $p = 2$ and 3 we have by Proposition 4.3,

$$N_p = \{z \in \mathbb{C}_p : |z - \mathrm{j}_p|_p > r_p\}.$$

For $p = 2$ and 3, put

$$\breve{\alpha}_p := \hat{\alpha}_p - \mathrm{j}_p \quad \text{and} \quad \breve{\beta}_p := \hat{\beta}_p - \mathrm{j}_p.$$

Note that for $p = 3$, we have

$$\breve{\alpha}_3(z) = \frac{(z^2 + 2 \cdot 3^2 z - 3^3)^2}{z} \quad \text{and} \quad \breve{\beta}_3(z) = \frac{(z^2 - 2 \cdot 3^5 z - 3^9)^2}{z^3}.$$

So, for $p = 2$ and 3 the rational map $\delta_p(z) := z^{-1}\breve{\beta}_p(z) - 1$ is a Laurent polynomial in $\frac{1}{z}\mathbb{Z}\big[\frac{1}{z}\big]$ satisfying $\|\delta_p\|_{r_p} \le 1$. In particular, for every $z$ in the set

$$\breve{N}_p := \{z' \in \mathbb{C}_p : |z'|_p > r_p\},$$

we have $|\breve{\beta}_p(z)|_p = |z|_p$, so $\breve{\beta}_p$ maps $\breve{N}_p$ into itself. By Lemma B.3 there is $\Delta_p(w)$ in $\frac{1}{w}\mathbb{Z}\big[\!\big[\frac{1}{w}\big]\!\big]$ such that $\|\Delta_p\|_{r_p} \le 1$ and such that the map

$$\begin{aligned} F_p \colon \breve{N}_p &\to \breve{N}_p \\ w &\mapsto F_p(w) := w(1 + \Delta_p(w)) \end{aligned}$$

is an inverse of $\breve{\beta}_p|_{\breve{N}_p}$.

We show below that $\boldsymbol{t}$ coincides with the map

$$\begin{aligned} \breve{\boldsymbol{t}} \colon N_p &\to \mathbb{C}_p \\ z &\mapsto \breve{\boldsymbol{t}}(z) := (\breve{\alpha}_p \circ F_p)(z - \mathrm{j}_p) + \mathrm{j}_p. \end{aligned}$$

Once this is established, the proposition follows from explicit computations using the estimates,

$$\|\Delta_p\|_{r_p} \le 1, \ \left\|\frac{\breve{\alpha}_2(w)}{w^2}\right\|_{2^{-4}} \le 1 \text{ for } p = 2, \quad \text{and} \quad \left\|\frac{\breve{\alpha}_3(w)}{w^3}\right\|_{3^{-3/2}} \le 1 \text{ for } p = 3.$$

By definition, for each $z$ in $\hat{N}_p$ the point $\check{\boldsymbol{t}}(z)$ is in the support of $T_p(z) = (\alpha_p)_* \circ \beta_p^*(z)$. Moreover, for every $z$ in $Y_{\mathrm{bad}}(\mathbb{C}_p)$ we have $|\check{\boldsymbol{t}}(z)|_p = |z|_p^p$, so by (5-1) and (5-3) we have $\check{\boldsymbol{t}}(z) = \boldsymbol{t}(z)$. Combined with Lemma B.4, this implies that $\check{\boldsymbol{t}}$ and $\boldsymbol{t}$ agree on $Y_{\mathrm{ord}}(\mathbb{C}_p) \cup Y_{\mathrm{bad}}(\mathbb{C}_p)$. In view of Proposition 4.3 and Lemma 4.6, to prove that $\check{\boldsymbol{t}}$ and $\boldsymbol{t}$ agree on $N_p \cap Y_{\mathrm{sups}}(\mathbb{C}_p)$ it is sufficient to show that for every $w$ in $\check{N}_p \cap \mathcal{M}_p$ we have $|(\check{\alpha}_p \circ F_p)(w)|_p \neq |w|_p^{1/p}$. Note that for every $w$ in $\check{N}_p$ we have $|F_p(w)|_p = |w|_p$. A direct computation shows that for $p = 2$ we have

$$|(\check{\alpha}_2 \circ F_2)(w)|_2 \begin{cases} = |w|_2^2 & \text{if } 2^{-4} < |w|_2 < 1; \\ \leq 2^{-8} & \text{if } |w|_2 = 2^{-4}; \\ = 2^{-12}/|w|_2 & \text{if } r_2 < |w|_2 < 2^{-4}, \end{cases}$$

and that for $p = 3$ we have

$$|(\check{\alpha}_3 \circ F_3)(w)|_3 \begin{cases} = |w|_3^3 & \text{if } 3^{-3/2} < |w|_3 < 1; \\ \leq 3^{-9/2} & \text{if } |w|_3 = 3^{-3/2}; \\ = 3^{-6}/|w|_3 & \text{if } r_3 < |w|_3 < 3^{-3/2}. \end{cases}$$

In all the cases we have $|(\check{\alpha}_p \circ F_p)(w)|_p \neq |w|_p^{1/p}$. This completes the proof of $\boldsymbol{t} = \check{\boldsymbol{t}}$, and of the proposition. □

*Proof of Theorem B.1.* We first prove (B-2), and the assertions about the Laurent series expansion. For $p = 2$ and 3, these are given by Proposition B.2. Assume $p \geq 5$. For each $e$ in $Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)$, let $\mathrm{j}_e$ be given by Proposition 4.3, and define $P_{\mathrm{sups}}(X) = \prod_{e \in Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)} (X - \mathrm{j}_e)$ as in the proof of this proposition. Since the reduction modulo $p$ of the polynomial $P_{\mathrm{sups}}$ is separable, for every $e$ in $Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)$ we have that $\mathrm{j}_e$ is in $\mathbb{Q}_p^{\mathrm{unr}}$. Put $\beta_e := \mathrm{j}_e$. Denote by $\hat{\boldsymbol{t}}$ the finite sum of Laurent series in the right-hand side of (3-3) for these choices of $(\beta_e)_{e \in Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)}$. It follows from Theorem 3.3 and Proposition 4.3 that $\hat{\boldsymbol{t}}$ converges on $N_p$, and by Lemma B.4 that for every $z$ in $Y_{\mathrm{bad}}(\mathbb{C}_p) \cup Y_{\mathrm{ord}}(\mathbb{C}_p)$ we have $\hat{\boldsymbol{t}}(z) = \boldsymbol{t}(z)$. We proceed to prove that for every $z$ in $\hat{N}_p := N_p \cap Y_{\mathrm{sups}}(\mathbb{C}_p)$ we also have $\hat{\boldsymbol{t}}(z) = \boldsymbol{t}(z)$.

Denote by $\Phi_p(X, Y)$ the modular polynomial of level $p$ defined in Section 2B. Note that for every $z$ in $Y_{\mathrm{bad}}(\mathbb{C}_p) \cup Y_{\mathrm{ord}}(\mathbb{C}_p)$ we have

$$\Phi_p(\hat{\boldsymbol{t}}(z), z) = \Phi_p(z, \hat{\boldsymbol{t}}(z)) = 0. \tag{B-4}$$

Since $\hat{\boldsymbol{t}}$ is analytic, (B-4) holds for every $z$ in $N_p$. In view of Lemma 4.6, this implies that for every $E$ in $\hat{N}_p$ we have either $v_p(\hat{\boldsymbol{t}}(E)) = \frac{1}{p} v_p(E)$, or

$$v_p(\hat{\boldsymbol{t}}(E)) \begin{cases} = p v_p(E) & \text{if } v_p(E) \in ]0, 1/(p+1)]; \\ \geq p v_p(E) & \text{if } v_p(E) = 1/p + 1; \\ = 1 - v_p(E) & \text{if } v_p(E) \in ]1/(p+1), p/(p+1)[. \end{cases} \tag{B-5}$$

We now prove that (B-5) holds for every $E$ in $\hat{N}_p$. Fix $e$ in $Y_{\mathrm{sups}}(\bar{\mathbb{F}}_p)$, and note that the function

$$\begin{aligned} \nu \colon ]0, p/(p+1)[ \cap \mathbb{Q} &\to \mathbb{Q} \\ r &\mapsto \nu(r) := \inf\{v_p(\hat{\boldsymbol{t}}(E)) : E \in \boldsymbol{D}(j(e)), v_p(E) = r\}, \end{aligned}$$

extends continuously to $]0,\ p/(p+1)[$. Thus, either (B-5) holds for every $E$ in $N_p \cap \boldsymbol{D}(j(e))$, or for every $E$ in this set we have $v_p(\hat{\boldsymbol{t}}(E)) = \frac{1}{p} v_p(E)$. So, to prove that (B-5) holds for every $E$ in $N_p \cap \boldsymbol{D}(j(e))$ it is sufficient to prove that it holds for some $E_0$ in $N_p \cap \boldsymbol{D}(j(e))$. Choose $E_0$ in $N_p \cap \boldsymbol{D}(j(e))$ such that $z_0 := j(E_0)$ satisfies

$$0 < \operatorname{ord}_p(z_0 - \mathrm{j}_e) < \frac{1}{p+1}.$$

By Theorem 3.3 we have

$$\operatorname{ord}_p(\hat{\boldsymbol{t}}(z_0) - z_0^p - pk(z_0)) \geq 1 - \operatorname{ord}_p(z_0 - \mathrm{j}_e) > \frac{p}{p+1}.$$

Since $\operatorname{ord}_p(z_0 - \mathrm{j}_e) < \frac{1}{p}$, we also have

$$\operatorname{ord}_p(\hat{\boldsymbol{t}}(z_0) - \mathrm{j}_e^p) = p \operatorname{ord}_p(z_0 - \mathrm{j}_e) < \frac{p}{p+1}.$$

Combined with $\operatorname{ord}_p(\mathrm{j}_e^p - \mathrm{j}_{e^{(p)}}) \geq 1$ and $\operatorname{ord}_p(pk(z_0)) \geq 1$, this implies

$$\operatorname{ord}_p(\hat{\boldsymbol{t}}(z_0) - \mathrm{j}_{e^{(p)}}) = p \operatorname{ord}_p(z_0 - \mathrm{j}_e), \tag{B-6}$$

and therefore (B-5) with $E = E_0$. This completes the proof that (B-5) holds for every $E$ in $\hat{N}_p$. In view of (B-4), Proposition 4.3, and Lemma 4.6, it follows that for every $z$ in $\hat{N}_p$ we have $\hat{\boldsymbol{t}}(z) = \boldsymbol{t}(z)$. By Theorem 3.3 we also obtain (B-2).

It remains to prove (B-3) for an arbitrary prime number $p$. Note that for $E$ in $Y_{\mathrm{ord}}(\mathbb{C}_p)$ this is given by Proposition 3.4 with $m = 1$, and that for $E$ in $Y_{\mathrm{bad}}(\mathbb{C}_p)$ this follows from the combination of (5-1), and of (5-3) with $n = p$. It remains to prove (B-3) for $E$ in $\hat{N}_p$. By the considerations above, and the proof of Proposition B.2, we have that (B-5) holds for every prime number $p$ and for every $E$ in $\hat{N}_p$. By Lemma 4.6 we deduce that:

(1) $\boldsymbol{t}$ maps

$$N_p' := \left\{ E \in Y(\mathbb{C}_p) : 0 < v_p(E) < \frac{1}{p+1} \right\}$$

onto $\hat{N}_p$, and for every $E$ in $\hat{N}_p$ the divisor $(\boldsymbol{t}|_{N_p'})^*(E)$ has degree $p$.

(2) $\boldsymbol{t}$ maps

$$S_p := \left\{ E \in Y(\mathbb{C}_p) : v_p(E) = \frac{1}{p+1} \right\}$$

onto $B_p := Y_{\mathrm{sups}}(\mathbb{C}_p) \setminus \hat{N}_p$, and for every $E$ in $B_p$ the divisor $(\boldsymbol{t}|_{S_p})^*(E)$ has degree $p+1$.

(3) $\boldsymbol{t}$ maps $A_p := \hat{N}_p \setminus (N_p' \cup S_p)$ onto itself, and for every $E$ in $A_p$ we have $(\boldsymbol{t}|_{A_p})^*(E) = [\boldsymbol{t}(E)]$.

The proof of (B-3) is divided in the following cases:

(1) For $E$ in $B_p$, we have $\boldsymbol{t}^*(E) = (\boldsymbol{t}|_{S_p})^*(E)$ and this divisor has degree $p+1$. Together with (B-4) this implies $T_p(E) = \boldsymbol{t}^*(E)$.

(2) For $E$ in $A_p$, we have $\boldsymbol{t}^*(E) = (\boldsymbol{t}|_{N'_p})^*(E) + (\boldsymbol{t}|_{A_p})^*(E)$ and this divisor has degree $p+1$. As in the previous case we conclude that $T_p(E) = \boldsymbol{t}^*(E)$.

(3) For $E$ in $N'_p \cup S_p$, we have $\boldsymbol{t}^*(E) = (\boldsymbol{t}|_{N'_p})^*(E)$ and this divisor is of degree $p$. Combined with (B-4) this implies that the divisor $T_p(E) - \boldsymbol{t}^*(E)$ has degree 1. On the other hand, by (B-5) the point $\boldsymbol{t}(E)$ is not in the support of $\boldsymbol{t}^*(E)$, so by (B-4) we have $T(E) - \boldsymbol{t}^*(E) = [\boldsymbol{t}(E)]$.

This completes the proof of (B-3), and of the theorem. □

## Acknowledgments

Menares and Rivera-Letelier thank Leon Takhtajan for references. Menares thanks Emmanuel Ullmo for sharing his interest in the questions we study here. He also acknowledges Rodolphe Richard for explaining him the basic ideas leading to Propositions 5.1 and 5.4. We thank the anonymous referees for their valuable comments that helped us improve the exposition.

During the preparation of this work the Herrero was supported by the Chilean CONICYT grant 21130412 and the Royal Swedish Academy of Sciences. Menares acknowledges partial support from FONDECYT grant 1171329. Rivera-Letelier acknowledges partial support from NSF grant DMS-1700291. The authors would like to thank the Pontificia Universidad Católica de Valparaíso, the University of Rochester and Universitat de Barcelona for hospitality during the preparation of this work.

## References

[Apostol 1976] T. M. Apostol, *Introduction to analytic number theory*, Springer, 1976. MR Zbl

[Baker and Rumely 2010] M. Baker and R. Rumely, *Potential theory and dynamics on the Berkovich projective line*, Math. Surv. Monogr. **159**, Amer. Math. Soc., Providence, RI, 2010. MR Zbl

[Berkovich 1990] V. G. Berkovich, *Spectral theory and analytic geometry over non-Archimedean fields*, Math. Surv. Monogr. **33**, Amer. Math. Soc., Providence, RI, 1990. MR Zbl

[Billingsley 1968] P. Billingsley, *Convergence of probability measures*, Wiley, New York, 1968. MR Zbl

[Bilu 1997] Y. Bilu, "Limit distribution of small points on algebraic tori", *Duke Math. J.* **89**:3 (1997), 465–476. MR Zbl

[Brink 2006] D. Brink, "New light on Hensel's lemma", *Expo. Math.* **24**:4 (2006), 291–306. MR Zbl

[Buzzard 2003] K. Buzzard, "Analytic continuation of overconvergent eigenforms", *J. Amer. Math. Soc.* **16**:1 (2003), 29–55. MR Zbl

[Clozel and Ullmo 2004] L. Clozel and E. Ullmo, "Équidistribution des points de Hecke", pp. 193–254 in *Contributions to automorphic forms, geometry, and number theory* (Baltimore, MD, 2002), edited by H. Hida et al., Johns Hopkins Univ. Press, Baltimore, MD, 2004. MR Zbl

[Clozel et al. 2001] L. Clozel, H. Oh, and E. Ullmo, "Hecke operators and equidistribution of Hecke points", *Invent. Math.* **144**:2 (2001), 327–351. MR Zbl

[Cohen 2007] H. Cohen, *Number theory, II: Analytic and modern tools*, Graduate Texts in Math. **240**, Springer, 2007. MR Zbl

[Coleman and McMurdy 2006] R. Coleman and K. McMurdy, "Fake CM and the stable model of $X_0(Np^3)$", *Doc. Math.* extra volume (2006), 261–300. MR Zbl

[Cox 2013] D. A. Cox, *Primes of the form $x^2 + ny^2$: Fermat, class field theory, and complex multiplication*, 2nd ed., Wiley, Hoboken, NJ, 2013. MR Zbl

[Deuring 1941] M. Deuring, "Die Typen der Multiplikatorenringe elliptischer Funktionenkörper", *Abh. Math. Sem. Hansischen Univ.* **14** (1941), 197–272. MR Zbl

[Diamond and Im 1995] F. Diamond and J. Im, "Modular forms and modular curves", pp. 39–133 in *Seminar on Fermat's last theorem* (Toronto, 1993-1994), edited by V. K. Murty, CMS Conf. Proc. **17**, Amer. Math. Soc., Providence, RI, 1995. MR Zbl

[Diamond and Shurman 2005] F. Diamond and J. Shurman, *A first course in modular forms*, Graduate Texts in Math. **228**, Springer, 2005. MR Zbl

[Dinh et al. 2020] T.-C. Dinh, L. Kaufmann, and H. Wu, "Dynamics of holomorphic correspondences on Riemann surfaces", *Int. J. Math.* **31**:5 (2020), 2050036, 21. MR Zbl

[Duke 1988] W. Duke, "Hyperbolic distribution problems and half-integral weight Maass forms", *Invent. Math.* **92**:1 (1988), 73–90. MR Zbl

[Dwork 1969] B. Dwork, "$p$-adic cycles", *Inst. Hautes Études Sci. Publ. Math.* **37** (1969), 27–115. MR Zbl

[Eskin and Oh 2006] A. Eskin and H. Oh, "Ergodic theoretic proof of equidistribution of Hecke points", *Ergodic Theory Dynam. Systems* **26**:1 (2006), 163–167. MR Zbl

[Fresnel and van der Put 2004] J. Fresnel and M. van der Put, *Rigid analytic geometry and its applications*, Progr. Math. **218**, Birkhäuser, Boston, 2004. MR Zbl

[Goren and Kassaei 2017] E. Z. Goren and P. L. Kassaei, "$p$-adic dynamics of Hecke operators on modular curves", preprint, 2017. arXiv

[Gross 1986] B. H. Gross, "On canonical and quasicanonical liftings", *Invent. Math.* **84**:2 (1986), 321–326. MR

[Hazewinkel 1978] M. Hazewinkel, *Formal groups and applications*, Pure Appl. Math. **78**, Academic Press, New York, 1978. MR Zbl

[Herrero et al. 2019] S. Herrero, R. Menares, and J. Rivera-Letelier, "$p$-adic distribution of CM points and Hecke orbits. II: Linnik equidistribution on the supersingular locus", preprint, 2019.

[Iwaniec 1987] H. Iwaniec, "Fourier coefficients of modular forms of half-integral weight", *Invent. Math.* **87**:2 (1987), 385–401. MR Zbl

[de Jong and Noot 1991] J. de Jong and R. Noot, "Jacobians with complex multiplication", pp. 177–192 in *Arithmetic algebraic geometry* (Texel, Netherlands, 1989), edited by G. van der Geer et al., Progr. Math. **89**, Birkhäuser, Boston, 1991. MR Zbl

[Kaneko and Zagier 1998] M. Kaneko and D. Zagier, "Supersingular $j$-invariants, hypergeometric series, and Atkin's orthogonal polynomials", pp. 97–126 in *Computational perspectives on number theory* (Chicago, 1995), edited by D. A. Buell and J. T. Teitelbaum, AMS/IP Stud. Adv. Math. **7**, Amer. Math. Soc., Providence, RI, 1998. MR Zbl

[Katz 1973] N. M. Katz, "$p$-adic properties of modular schemes and modular forms", pp. 69–190 in *Modular functions of one variable, III* (Antwerp, Belgium, 1972), edited by W. Kuyk and J.-P. Serre, Lecture Notes in Math. **350**, Springer, 1973. MR Zbl

[Kohnen 2003] W. Kohnen, "Transcendence of zeros of Eisenstein series and other modular functions", *Comment. Math. Univ. St. Pauli* **52**:1 (2003), 55–57. MR Zbl

[Lang 1973] S. Lang, *Elliptic functions*, Addison-Wesley, Reading, MA, 1973. MR Zbl

[Lang 1994] S. Lang, *Algebraic number theory*, 2nd ed., Graduate Texts in Math. **110**, Springer, 1994. MR Zbl

[Linnik 1968] Y. V. Linnik, *Ergodic properties of algebraic fields*, Ergebnisse der Mathematik **45**, Springer, 1968. MR Zbl

[Linnik and Skubenko 1964] Y. V. Linnik and B. F. Skubenko, "Asymptotic distribution of integral matrices of third order", *Vestnik Leningrad. Univ. Ser. Mat. Meh. Astronom.* **19**:3 (1964), 25–36. In Russian. MR Zbl

[Mestre 1986] J.-F. Mestre, "La méthode des graphes: exemples et applications", pp. 217–242 in *Proc. Int. Conf. on Class Numbers and Fundamental Units of Algebraic Number Fields* (Katata, Japan, 1986), edited by Y. Yamamoto and H. Yokoi, Nagoya Univ., 1986. MR Zbl

[Michel and Venkatesh 2006] P. Michel and A. Venkatesh, "Equidistribution, $L$-functions and ergodic theory: on some problems of Yu. Linnik", pp. 421–457 in *Proc. Int. Congr. Math., II* (Madrid, 2006), edited by M. Sanz-Solé et al., Eur. Math. Soc., Zürich, 2006. MR Zbl

[Richard 2018] R. Richard, "Répartition galoisienne ultramétrique d'une classe d'isogénie de courbes elliptiques: le cas de la mauvaise réduction", *J. Théor. Nombres Bordeaux* **30**:1 (2018), 1–18. MR Zbl

[Rivera-Letelier 2003] J. Rivera-Letelier, "Dynamique des fonctions rationnelles sur des corps locaux", pp. 147–230 in *Geometric methods in dynamics, II*, edited by W. de Melo et al., Astérisque **287**, Soc. Math. France, Paris, 2003. MR Zbl

[Roquette 1970] P. Roquette, *Analytic theory of elliptic functions over local fields*, Hamburger Math. Einzelschriften (N.F.) **1**, Vandenhoeck & Ruprecht, Göttingen, 1970. MR Zbl

[Shimura 1971] G. Shimura, *Introduction to the arithmetic theory of automorphic functions*, Publ. Math. Soc. Japan **11**, Iwanami Shoten, Tokyo, 1971. MR Zbl

[Siegel 1935] C. L. Siegel, "Über die Classenzahl quadratischer Zahlkörper", *Acta Arith.* **1**:1 (1935), 83–86. Zbl

[Silverman 1994] J. H. Silverman, *Advanced topics in the arithmetic of elliptic curves*, Graduate Texts in Math. **151**, Springer, 1994. MR Zbl

[Silverman 2009] J. H. Silverman, *The arithmetic of elliptic curves*, 2nd ed., Graduate Texts in Math. **106**, Springer, 2009. MR Zbl

[Tate 1995] J. Tate, "A review of non-Archimedean elliptic functions", pp. 162–184 in *Elliptic curves, modular forms, & Fermat's last theorem* (Hong Kong, 1993), edited by J. Coates and S.-T. Yau, Ser. Number Theory **1**, Int. Press, Cambridge, MA, 1995. MR Zbl

[Zhang 2001] S. Zhang, "Heights of Heegner points on Shimura curves", *Ann. of Math.* (2) **153**:1 (2001), 27–147. MR Zbl

Communicated by Shou-Wu Zhang
Received 2018-11-03 Revised 2019-11-16 Accepted 2020-02-06

sebastian.herrero.m@gmail.com *Instituto de Matemáticas, Pontificia Universidad Católica de Valparaíso, Chile*

rmenares@mat.uc.cl *Facultad de Matemáticas, Pontificia Universidad Católica de Chile, Santiago, Chile*

riveraletelier@gmail.com *Mathematics Department, University of Rochester, NY, United States*

ALGEBRA AND NUMBER THEORY 14:5 (2020)
https://doi.org/10.2140/ant.2020.14.1291

# Roots of $L$-functions of characters over function fields, generic linear independence and biases

Corentin Perret-Gentil

We first show joint uniform distribution of values of Kloosterman sums or Birch sums among all extensions of a finite field $\mathbb{F}_q$, for almost all couples of arguments in $\mathbb{F}_q^\times$, as well as lower bounds on differences. Using similar ideas, we then study the biases in the distribution of generalized angles of Gaussian primes over function fields and primes in short intervals over function fields, following recent works of Rudnick and Waxman, and Keating and Rudnick, building on cohomological interpretations and determinations of monodromy groups by Katz. Our results are based on generic linear independence of Frobenius eigenvalues of $\ell$-adic representations, that we obtain from integral monodromy information via the strategy of Kowalski, which combines his large sieve for Frobenius with a method of Girstmair. An extension of the large sieve is given to handle wild ramification of sheaves on varieties.



## 1. Introduction and statement of the results

Throughout, $p$ will denote a prime larger than 5 and $q$ a power of $p$.

**1A.** ***Kloosterman and Birch sums.*** For an integer $n \geq 1$, and $a \in \mathbb{F}_{q^n}^\times$, we consider the Kloosterman sums

$$\mathrm{Kl}_{r,q^n}(a) = \frac{1}{q^{n(r-1)/2}} \sum_{\substack{x_1,\dots,x_r \in \mathbb{F}_{q^n}^\times \\ x_1 \dots x_r = a}} e\left(\frac{\mathrm{tr}(x_1 + \cdots + x_r)}{p}\right) \tag{1}$$

of integer rank $r \geq 2$, as well as the Birch sums

$$\mathrm{Bi}_{q^n}(a) = \frac{1}{q^{n/2}} \sum_{x \in \mathbb{F}_{q^n}^\times} e\left(\frac{\mathrm{tr}(ax + x^3)}{p}\right). \tag{2}$$

*MSC2010:* primary 14G10; secondary 11J72, 11N36, 11R58, 11T23.
*Keywords:* exponential sums, linear independence, L-functions, large sieve, characters, function fields, Kloosterman sums.

Here, we adopt the usual notation $e(z) = \exp(2\pi i z)$ for any $z \in \mathbb{C}$, and $\mathrm{tr} : \mathbb{F}_{q^n} \to \mathbb{F}_p$ is the field trace.

For convenience, let us define the rank of $\mathrm{Bi}_{q^n}$ to be $r = 2$, and for $r \geq 2$, we let

$$f_{q^n} = \mathrm{Kl}_{r,q^n} \ (r \geq 2) \qquad \text{or} \qquad \mathrm{Bi}_{q^n} \ (r = 2) \tag{3}$$

for every integer $n \geq 1$. By the Deligne–Katz equidistribution theorem [Katz 1988] for Kloosterman sums and Livné's work [1987] for Birch sums (see also [Katz 1990]), as $q^n \to \infty$ the values

$$\{f_{q^n}(a) : a \in \mathbb{F}_{q^n}^\times\}$$

equidistribute in

$$\Omega_r = \begin{cases} [-r, r] \subset \mathbb{R} & \text{for } r \text{ even}, \\ \{z \in \mathbb{C} : |z| \leq r\} & \text{for } r \text{ odd}, \end{cases}$$

with respect to the pushforward $\mathrm{tr}_* \mu_r$ of the Haar measure $\mu_r$ on the compact group

$$G_r(\mathbb{C}), \qquad \text{where } G_r := \begin{cases} \mathrm{SU}_r & \text{for } r \text{ odd}, \\ \mathrm{USp}_r & \text{for } r \text{ even} \end{cases}$$

(e.g., the Sato–Tate measure when $r = 2$). These statements encompass bounds on $f_{q^n}$ (e.g., Deligne's bound for hyper-Kloosterman sums), and the fact that $f_{q^n}$ is real-valued whenever $r$ is even. Moreover, they can alternatively be phrased as properties of the "angles" of Kloosterman and Birch sums, i.e., the

$$\theta_{1,f,q}(x), \dots, \theta_{r,f,q}(x) \in [0, 1],$$

such that

$$f_{q^n}(x) = \sum_{i=1}^{r} e(n\theta_{i,f,q}(x)) \qquad \text{for all } n \geq 1, \ x \in \mathbb{F}_q^\times \tag{4}$$

(whose existence follows from profound work of Grothendieck, Deligne, Katz and others, and will be recalled in due time): they are distributed like the eigenvalues of a Haar-random matrix in $G_r(\mathbb{C})$.

Our first main result is the following generic linear independence statement:

**Theorem 1.1** (generic pairwise linear independence). *For $r \geq 2$ fixed, let $f$ be as in* (3)*, and let*

$$E_r := \dim G_r + \tfrac{1}{2} \operatorname{rank} G_r = \begin{cases} \frac{1}{2}(2r^2 + r - 3) & \text{for } r \text{ odd}, \\ \frac{1}{4}(2r^2 + 3r) & \text{for } r \text{ even}. \end{cases}$$

*For almost all $a, b \in \mathbb{F}_q^\times$, that is for*

$$(q-1)^2 \left[1 + O_{r,p}\left(\frac{\log q}{q^{1/(2E_r)}}\right)\right] = (q-1)^2(1 + o_{r,p}(1)) \tag{5}$$

*of them, the angles*

$$1, \quad \theta_{j,f,q}(a), \quad \theta_{j,f,q}(b) \quad \text{with} \quad \begin{cases} 1 \leq j \leq r-1 & \text{for } r \text{ odd}, \\ 1 \leq j \leq r/2 & \text{for } r \text{ even}, \end{cases} \tag{6}$$

*are $\mathbb{Q}$-linearly independent. The implied constants depend only on $r$, $p$, and only on $r$ in the case of Kloosterman sums.*

$$\mathbb{F}_p^\times \text{ (fixed)} \longrightarrow \mathbb{F}_q^\times \longrightarrow \mathbb{F}_{q^n}^\times \longrightarrow \bigcup_{m\geq 1} \mathbb{F}_{q^m}^\times = \overline{\mathbb{F}}_q^\times$$
$$\ni$$
$$a, b$$

**Figure 1.** The asymptotic setting for Section 1A.

**Remark 1.2.** The restriction on $j$ in (6) is necessary since $\sum_{j=1}^r \theta_{j,f,q}(x) = 0$, and if $r$ is even, the angles come by pairs: $\theta_{r/2+j,f,q}(x) = -\theta_{j,f,q}(x)$ $(1 \leq j \leq r/2)$.

**Remark 1.3.** Actually, we will more generally prove Theorem 1.1 for almost all tuples of $t \geq 1$ arguments, when $t = o(\sqrt{\log q})$ (e.g., $t$ fixed), with (5) replaced by

$$(q-1)^t \left(1 + O_{r,p}\left(\frac{(r^{\delta_{r \text{ odd}}} C)^t \log q}{q^{1/(tE_r)}}\right)\right) \tag{7}$$

for an absolute constant $C \geq 1$. The implied constants depends again only on $r$ in the case of Kloosterman sums.[1]

This has several interesting consequences. First, we obtain the *joint distribution* of almost all pairs of values of $f$ in extensions of a fixed base field:

**Corollary 1.4.** *For $r \geq 2$, let $f$ be as in* (3)*, a Kloosterman or Birch sum. For all but*

$$O_{r,p}\big((q-1)^2 (\log q) q^{-1/(2E_r)}\big)$$

*couples $a, b \in \mathbb{F}_q^\times$, the random vector*

$$X_{a,b} = \big((f_{q^n}(a),\, f_{q^n}(b))\big)_{1\leq n\leq N}$$

(*with the uniform measure on* $[1, N] \cap \mathbb{N}$) *converges in law as $N \to \infty$ to*

$$(\mathrm{tr}(g_1), \mathrm{tr}(g_2)),$$

*with $g_1, g_2$ independent uniformly distributed in a maximal torus of $G_r(\mathbb{C})$. Explicitly,* $\mathrm{tr}(g_i)$ *is distributed like*

$$\begin{cases} \sum_{j=1}^{r/2} 2\cos(2\pi\theta_j) & \text{for } r \text{ even}, \\ \sum_{j=1}^{r-1} e(\theta_j) + e\Big(-\sum_{j=1}^{r-1} \theta_j\Big) & \text{for } r \text{ odd}, \end{cases} \tag{8}$$

*with $\theta_j$ independent uniform in* $[0, 1]$. *Equivalently, the distribution of* $\mathrm{tr}(g_i)$ *is that of* $\mathrm{tr}(h_i^m)$ *for any $m \geq r$ and $h_i$ uniform in $G_r(\mathbb{C})$ with respect to the Haar measure. The implied constant in Landau's notation depends only on $r$ in the case of Kloosterman sums.*

[1] Here and from now on, $\delta_B$ will denote the Kronecker symbol with respect to a binary variable $B$, i.e., $\delta_B = 1$ if $B$ is true, 0 otherwise. In particular, $r^{\delta_{r \text{ odd}}}$ is equal to $r$ if the latter is odd, and to 1 otherwise.

**Remark 1.5.** Applying Deligne's equidistribution theorem and [Katz 1988; 1990] would show that $\big(f_{q^n}(a+b_1),\dots,f_{q^n}(a+b_t)\big)_{a\in\mathbb{F}_{q^n},\ a+b_i\neq 0}$ converges in law (with respect to the uniform measure), as $q^n\to\infty$, to a random vector in $\Omega_r^t$ distributed with respect to the product measure $(\mathrm{tr}_*\,\mu_r)^{\otimes t}$, when $b_i\in\mathbb{F}_{q^n}$ are $t=o(\log(q^n))$ distinct shifts (see, e.g., [Perret-Gentil 2017], where the dependencies of the errors from [Fouvry et al. 2015] with respect to $t$ are made explicit). However, this only gives information among values that are explicitly related, by fixed shifts.

**Remark 1.6** (discrepancy). For the distribution of a single Kloosterman sum of rank 2, conditionally on a linear independence hypothesis, Ahmadi and Shparlinski [2010] also obtained bounds on the discrepancy, using lower bounds arising from Baker's theorem. Their results are stated for curves, but the last paragraph of [Ahmadi and Shparlinski 2010, Section 5.2] explains how they readily extend to Kloosterman sums. Our Theorem 1.1 shows that their discrepancy bounds hold for almost all arguments, and using the same technique, a bound on the discrepancy in Corollary 1.4 could as well be given.

Another corollary is the following absence of bias among values of Birch sums and Kloosterman sums in extensions:

**Corollary 1.7.** *Let $f_{q^n}$ be either* $\mathrm{Kl}_{r,q^n}$ *($r\geq 2$ even),* $\mathrm{Bi}_{q^n}$*, or — if $r\geq 3$ is odd —* $\mathrm{Re}\,\mathrm{Kl}_{r,q^n}$ *or* $\mathrm{Im}\,\mathrm{Kl}_{r,q^n}$*. For all but $O_{r,p}\big((q-1)^2(\log q)q^{-1/(2E_r)}\big)$ couples $a,b\in\mathbb{F}_q^\times$, we have*

$$\mathbb{P}_{n\leq N}(f_{q^n}(a)<f_{q^n}(b)):=\frac{|\{1\leq n\leq N: f_{q^n}(a)<f_{q^n}(b)\}|}{N}\to 1/2\qquad as\quad N\to\infty.$$

*The implied constant in Landau's notation depends only on $r$ in the case of Kloosterman sums.*

Finally, Theorem 1.1 also yields the following lower bounds, through the method of Bombieri and Katz [2010]. The first is not explicit and the value of $n$ is not effective, while the second is weaker but does not suffer from these issues.

**Corollary 1.8.** *For $r\geq 2$, let $f$ be as in* (3)*. For every $\varepsilon>0$ and all but $O_{r,p}\big((q-1)^2(\log q)q^{-1/(2E_r)}\big)$ couples $a,b\in\mathbb{F}_q^\times$, we have*:

(1) *For every $n$ large enough* (*with respect to $q,r,\varepsilon,a,b$*),

$$|f_{q^n}(a)-f_{q^n}(b)|\geq q^{-\varepsilon n(r-1)}.$$

(2) *When $r=2$, for every $n\geq 1$ large enough with respect to $p$,*

$$|f_{q^n}(a)-f_{q^n}(b)|\geq (2/\pi^2)\begin{cases} q^{-2^{26}3^3\pi p^3\log(4p)\log(2n+1/2)}\\ q^{-C_p\log\left(\frac{n}{e}+\frac{2n+1/2}{q}\right)\frac{\log q}{\max(\log q,2)}}\end{cases}$$

*with* $C_p=1175\big(5.205+0.946\log\big(\frac{1}{2}(p-1)\big)\big)(p-1)^4$.

**Remark 1.9.** The second bound in (2) uses Gouillon's improvement [2006] on the Baker–Wüstholz theorem [1993] instead of the latter. The condition on $n$ is only to simplify the expression above: the

bound in the proof is fully explicit. Moreover, the first inequality in (2) is valid for any $n \geq 1$. We can also update the lower bound of [Bombieri and Katz 2010, Corollary 4.3(ii)] to (assuming $p \geq 5$):

$$|\operatorname{Kl}_{r,p^n}(a)| \geq (2/\pi)q^{-2C_p \log\left(\frac{n}{e}+\frac{4n+1}{q}\right)\frac{\log q}{\max(\log q,2)}},$$

with $C_p$ as above.

**1B.** ***Angles of Gaussian primes over function fields.*** Recently, Rudnick and Waxman [2019] studied refined statistics of angles of Gaussian primes $p = a + ib \in \mathbb{Z}[i]$, after Hecke's equidistribution result and the works that ensued. To give motivation for a conjecture they proposed, they developed a function field model where an analogue holds unconditionally.

Explicitly (see [Rudnick and Waxman 2019, Section 1.3, Section 6]), consider the quadratic extension $\mathbb{F}_q(S)$ of the function field $\mathbb{F}_p(T)$, $S = \sqrt{-T}$, with the norm $N(f(S)) = f(S)f(-S)$. The analogue of the unit circle is

$$\mathbb{S}_q^1 := \{u \in \mathbb{F}_q[[S]]^\times : u(0) = 1,\ N(u) = 1\},$$

and we have a well-defined map $U : \mathbb{F}_q[S]\backslash\{0\} \to \mathbb{S}_q^1,\ f \mapsto f/\sqrt{N(f)}$, that actually only depends on the ideal $(f)$. For an integer $k \geq 1$, the "circle" $\mathbb{S}_q^1$ can be divided into $q^\kappa$ sectors ($\kappa = \lfloor k/2 \rfloor$), $\operatorname{Sec}(u,k) := \{v \in \mathbb{S}_q^1 : v \equiv u \pmod{S^k}\}$, which are parametrized by

$$u \in \mathbb{S}_{k,q}^1 := \{u \in R_{k,q} : u(0) = 1,\ N(u) = 1\}, \quad R_{k,q} := (\mathbb{F}_q[S]/(S^k))^\times. \tag{9}$$

Rudnick and Waxman started by showing that if $k \leq n$ and

$$N_{k,n}(u) := |\{\mathfrak{p} \trianglelefteq \mathbb{F}_q[S] \text{ prime} : \deg(\mathfrak{p}) = n,\ U(\mathfrak{p}) \in \operatorname{Sec}(u,k)\}|$$

is the number of primes of fixed degree lying in a sector given by $u \in \mathbb{S}_{k,q}^1$, then there is equidistribution in the sectors whenever $\kappa < n/2$:

$$N_{k,n}(u) = \frac{|\{\mathfrak{p} \trianglelefteq \mathbb{F}_q[S] \text{ prime} : \deg(\mathfrak{p}) = n\}|}{|\mathbb{S}_{k,q}^1|} + O(q^{n/2}) = \frac{q^n/n}{q^\kappa} + O(q^{n/2}),$$

with an absolute implied constant.[2] Using a deep result of Katz [2017] (based on Deligne's equidistribution theorem and the computation of a monodromy group), they then got an unconditional analogue [Rudnick and Waxman 2019, Theorem 1.3] of their conjecture for $\mathbb{Z}[i]$ [Rudnick and Waxman 2019, Conjecture 1.2] on the variance of $N_{k,n}$ among all sectors.

The notion of Chebyshev bias for primes in arithmetic progressions, studied in depth by Rubinstein and Sarnak [1994], was extended to function fields by Cha [2008]. Further cases of biases in function fields have been considered recently [Cha and Kim 2010; Cha et al. 2016; 2017; Devin and Meng 2018], particularly in families of curves.

[2]The dependencies of the error with respect to $k$ are not explicit in [Rudnick and Waxman 2019], but keeping track of them during the arguments shows that the error in the expression for $N_{k,n}(u)$ above is $O\big(q^{n/2}\kappa/n + \tau(n)^{1/2}q^{n/2}/n\big)$ (recall that we assume that $p \geq 7$), where $\tau$ is the number of divisors function.

Similarly, one may ask whether there is a bias in the distribution of prime ideals among different sectors as above. To do so, for $u_1, \dots, u_R \in \mathbb{S}^1_{k,q}$ distinct, we may look at the $\mathbb{R}^R$-valued random vector

$$X_{k,N}(\boldsymbol{u}) := \big(X_{k,N}(u_1), \dots, X_{k,N}(u_R)\big), \quad \text{where } X_{k,N}(u_r) := \left(\frac{q^\kappa n}{q^{n/2}}\left(N_{k,n}(u_r) - \frac{q^n/n}{q^\kappa}\right)\right)_{1\le n\le N}$$

(with the uniform measure on $[1, N]\cap\mathbb{N}$). The normalization is chosen so that $X_{k,N}(u_r)$ is bounded as $N \to \infty$ (with $q, k$ fixed), which will be clear later on.

We recall that key inputs in [Rubinstein and Sarnak 1994; Cha 2008] to study biases finely are hypotheses about linear independence of roots of $L$-functions, also known as grand simplicity hypotheses (GSH). These are very strong statements and wide open conjectures.

Our second main result is a generic linear independence statement in the setting above, in the same spirit as Theorem 1.1. It concerns roots

$$e(\pm\theta_{\Xi,j}) \quad (1 \le j \le d'(\Xi)), \quad \theta_{\Xi,j} \in [0, 1], \tag{10}$$

of (normalized) $L$-functions associated to characters $\Xi$ of $\mathbb{S}^1_{k,q}$ with conductor $3 \le d(\Xi) \le 2\kappa - 1$, where $d'(\Xi) := (d(\Xi) - 1)/2$ (these will be defined more precisely in Section 3). The analogue of GSH is:

**Hypothesis 1.10.** The angles $\theta_{\Xi,j}$, for $\Xi \in \widehat{\mathbb{S}}^1_{k,q}$, $1 \le j \le d'(\Xi)$, are $\mathbb{Q}$-linearly independent.

Towards Hypothesis 1.10, we show:

**Theorem 1.11** (generic linear independence). *Assume that $p > k$ and let $t = o(\log|\mathbb{S}^1_{k,q}|)$ (e.g., $t$ fixed). For almost all subsets $S \subset \widehat{\mathbb{S}}^1_{k,q}$ of size $t$, that is for*

$$\binom{q^\kappa}{t}\left(1 + O_{k,p}\left(\frac{C_{k,p}^t \log q}{q^{1/(2t(2\kappa^2-3\kappa+1))}}\right)\right) = \binom{q^\kappa}{t}(1 + o_{k,p}(1))$$

*of them, with $C_{k,p} \ge 1$ depending only on $k$, $p$, the elements*

$$1, \quad \theta_{\Xi,j} \qquad (\Xi \in S,\ 1 \le j \le d'(\Xi))$$

*are $\mathbb{Q}$-linearly independent.*

**Remark 1.12.** Hypothesis 1.10 would be Theorem 1.11 with $S = \mathbb{S}^1_{k,q}$. This is a very strong statement, whose validity may be delicate depending on the relative size of the parameters. Indeed, unlike in the number field situation, there are examples of families of $L$-functions over function fields where linear independence is not satisfied (although with $q$ fixed, and eventually growing genus); see, e.g., [Kowalski 2008b, Section 6; Cha 2008, Section 5; Li 2018].

**Remark 1.13.** One can get the explicit dependency of the base $C_{k,p}$ with respect to $k$, $p$ in Theorem 1.11, at the cost of a weaker error, replacing the latter by

$$O_{k,p}\left(\frac{(C(k+1)^{k+1})^t \log\log q}{\log q}\right)$$

with $C$ absolute. Under a group theoretic conjecture, one could do so while keeping the strength of Theorem 1.11; see Remark 5.18.

Let us now explain how this relates to biases and the random vectors $X_{k,N}(\boldsymbol{u})$ defined above. We adapt classical arguments [Rubinstein and Sarnak 1994; Martin and Ng 2017; Devin 2019] to the function field setting, as in [Cha 2008; Devin and Meng 2018], to show:

**Theorem 1.14** (limiting distribution, expected value). *The random vector $X_{k,N}(\boldsymbol{u})$ admits a compactly supported limiting distribution as $N \to \infty$ with $\kappa < N/2$ fixed. Namely, it converges in law to a $\mathbb{R}^R$-valued random variable $X_k(\boldsymbol{u})$. Moreover, the expected value of the latter is*

$$\mathbb{E}(X_k(\boldsymbol{u})) = \big(-\big|\{b \in \mathbb{S}^1_{k,q} : b^2 = u_r\}\big|/2\big)_{1\le r\le R} \subset \big\{-\tfrac{1}{2}, 0\big\}^R,$$

*which means that there should be a bias towards sectors parametrized by nonsquares.*

**Theorem 1.15** (continuity, symmetry, bias). *If Hypothesis 1.10 holds and $R < \frac{1}{2}(\kappa - 1)$ is an integer, the distribution of $X_k(\boldsymbol{u})$ is*

(1) *absolutely continuous: there exists a Lebesgue integrable function $f$ on $\mathbb{R}^R$ such that $\mathbb{P}(X_k(\boldsymbol{u}) \in A) = \int_A f\, d\boldsymbol{x}$ for all Borel subsets $A \subset \mathbb{R}^R$;*

(2) *symmetric with exchangeable components around its mean: for $X_k^0(\boldsymbol{u}) := X_k(\boldsymbol{u}) - \mathbb{E}(X_k(\boldsymbol{u}))$, we have*

$$X_k^0(\boldsymbol{u}) \sim -X_k^0(\boldsymbol{u}),\ \sigma(X_k^0(\boldsymbol{u}))$$

*for any permutation $\sigma \in \mathfrak{S}_R$ of the coordinates.*

*Hence,*

$$\lim_{N\to\infty} \mathbb{P}\big(X_{k,N}(u_1) < \cdots < X_{k,N}(u_R)\big) = \mathbb{P}\big(X_k(\boldsymbol{u})_1 < \cdots < X_k(\boldsymbol{u})_R\big),$$

*which is $1/R!$ if the $u_i$ are all squares or all nonsquares. If $u_2$ is a square while $u_1$ is not, and $\kappa > 5$, then $\lim_{N\to\infty} \mathbb{P}\big(X_{k,N}(u_1) < X_{k,N}(u_2)\big) < \frac{1}{2}$.*

**Remark 1.16.** The restriction $R < \frac{1}{2}(\kappa - 1)$, rather strong with respect to the maximum $R = q^\kappa$, comes from the fact that the $L$-functions have finitely many zeros, in contrast with the number field case.

Hence, our generic linear independence statement, Theorem 1.11, implies the following towards an unconditional Theorem 1.15:

**Corollary 1.17** (of Theorem 1.11). *Assuming that $p > k$, the limiting distribution $X_k(\boldsymbol{u})$ of Theorem 1.14 is*

(1) *continuous: $\mathbb{P}(X_k(\boldsymbol{u}) = \boldsymbol{a}) = 0$ for any $\boldsymbol{a} \in \mathbb{R}^R$. In particular, for $u \in \mathbb{S}^1_{k,q}$, $\lim_{N\to\infty} \mathbb{P}(X_{k,N}(u) > 0) = \mathbb{P}(X_k(u) > 0)$;*

(2) *a pushforward of the Lebesgue measure on a torus of dimension*

$$\gg_{\varepsilon,k} (\log|\mathbb{S}^1_{k,q}|)^{1-\varepsilon}, \quad \text{for any } \varepsilon > 0.$$

**Remark 1.18.** Concerning the stronger properties of Theorem 1.15 (absolute continuity, symmetry), Devin [2019] and Martin and Ng [2017] have shown that they hold under weaker conditions than full linear independence. However, we cannot exploit these here since their statements always involve all the roots/eigenvalues, while results obtained from the large sieve will be limited to a small subset.

**1C.** ***Prime polynomials in short intervals.*** Some of the techniques in [Rudnick and Waxman 2019] actually originate from [Keating and Rudnick 2014], which showed function field analogues of a conditional result of Goldston–Montgomery on primes in short intervals and of a conjecture of Hooley on the variance of primes in arithmetic progressions with fixed modulus.

For $A \in \mathbb{F}_q[T]$ of degree $n \geq 1$ and $1 \leq h \leq n$,

$$\nu_h(A) := \sum_{\substack{f \in \mathbb{F}_q[T] \\ \deg(f-A) \leq h}} \Lambda(f)$$

counts prime polynomials in a "short interval" around $A$, weighted by the function field von Mangoldt function $\Lambda$ (defined by $\Lambda(f) = \deg(P)$ if $f = P^k$, $P \in \mathbb{F}_q[T]$ prime, $\Lambda(f) = 0$ otherwise). The mean value over the centers $A$ having degree $n$ is

$$\mathbb{E}_{q,n}(\nu_h) := \frac{1}{q^n} \sum_{\substack{A \in \mathbb{F}_q[T] \text{ monic} \\ \deg(A)=n}} \nu_h(A) = q^{h+1}\left(1 - \frac{1}{q^n}\right) \tag{11}$$

(see [Keating and Rudnick 2014, (2.7), Lemma 4.3]). Keating and Rudnick, [2014, Theorem 2.1], using another equidistribution result of Katz [2013b] when $h < n-3$, computed the corresponding variance explicitly, obtaining an unconditional analogue of the Goldston–Montgomery result mentioned above.

Any monic $A \in \mathbb{F}_q[T]$ of degree $n$ can be written uniquely as

$$A = T^{h+1}B + C \quad \text{with} \quad \begin{cases} B \text{ monic, } \deg(B) = n-h-1 \\ \deg(C) \leq h, \end{cases}$$

and $\nu_h(A) = \nu_h(T^{h+1}B)$ only depends on $B$. This observation allows us to fix $n-h =: m$ and take $n \to \infty$. For $B_1, \ldots, B_R \in \mathbb{F}_q[T]$ distinct and monic of degree $m-1$, we can study the $\mathbb{R}^R$-valued random vector of biases

$$X_{m,N}(\boldsymbol{B}) := \big(X_{m,N}(B_1), \ldots, X_{m,N}(B_R)\big),$$

where

$$X_{m,N}(B_r) := \left(\frac{q^m}{q^{n/2+1}}\big(\nu_{n-m}(T^{n-m+1}B_r) - \mathbb{E}_{q,n}(\nu_{n-m})\big)\right)_{1 \leq n \leq N}$$

(with the uniform measure on $[1, N] \cap \mathbb{N}$), the expected values being those in (11). Again, the normalization is chosen so that $X_{m,N}(u_r)$ is bounded as $N \to \infty$ (with $q, m$ fixed), which will be clear later on.

In this setting, we obtain results analogous to those exposed in Section 1B. Let

$$e(\theta_{\chi,j}) \quad (1 \leq j \leq d-1), \quad \theta_{\chi,j} \in [0,1], \tag{12}$$

be the roots associated to the $L$-function associated to an even Dirichlet character $\chi$ modulo $T^m \in \mathbb{F}_q[T]$ (see Section 3 for the precise definitions), for $2 \leq d \leq m$.

**Hypothesis 1.19.** The angles $\theta_{\chi,j}$, for $\chi \pmod{T^m}$ even, $1 \leq j \leq \operatorname{cond}(\chi)-2$, are $\mathbb{Q}$-linearly independent.

**Theorem 1.20** (generic linear independence). *Assume that $m$ is odd, $p > m$ and $t = o(\log(q^{m-1}))$ (e.g., $t$ fixed). For almost all subsets $S$ of size $t$ of even Dirichlet characters mod $T^m$, that is for*

$$\binom{q^{m-1}}{t}\left(1 + O_{p,m}\left(\frac{C_{m,p}^t \log q}{q^{1/(2t(m-2)^2)}}\right)\right) = \binom{q^{m-1}}{t}(1 + o_{p,m}(1))$$

*of them, with $C_{m,p} \geq 1$ depending only on $p, m$, the elements*

$$1, \quad \theta_{\chi,j} \qquad (\chi \in S,\ 1 \leq j \leq \operatorname{cond}(\chi) - 2)$$

*are $\mathbb{Q}$-linearly independent.*

**Theorem 1.21** (limiting distribution, expected value). *The random vector $X_{m,N}(\boldsymbol{B})$ admits a compactly supported limiting distribution as $N \to \infty$ with $m > 3$ fixed. Namely, it converges in law to a $\mathbb{R}^R$-valued random variable $X_m(\boldsymbol{B})$. Moreover, the latter has mean zero.*

**Remark 1.22.** There is no bias here, unlike in Theorem 1.14, simply because the von Mangoldt weight was kept.

**Theorem 1.23** (continuity, symmetry). *If Hypothesis 1.19 holds and $R < m/2 - 1$, the distribution of $X_m(\boldsymbol{B})$ is absolutely continuous, and symmetric with exchangeable components. In particular,*

$$\lim_{N\to\infty} \mathbb{P}\big(X_{m,N}(B_1) < \cdots < X_{m,N}(B_R)\big) = \frac{1}{R!}.$$

Towards an unconditional Theorem 1.23, we obtain:

**Corollary 1.24** (of Theorem 1.21). *Assuming $m$ odd and $p > m$, the limiting distribution $X_m(\boldsymbol{B})$ from Theorem 1.21 is*

(1) *continuous*: $\mathbb{P}(X_m(\boldsymbol{B}) = \boldsymbol{a}) = 0$ *for any $\boldsymbol{a} \in \mathbb{R}^R$. In particular, for $B \in \mathbb{F}_q[T]$ of degree $m-1$,*

$$\lim_{N\to\infty} \mathbb{P}(X_{m,N}(B) > 0) = \mathbb{P}(X_m(B) > 0);$$

(2) *a pushforward of the Lebesgue measure on a torus of dimension*

$$\gg_{\varepsilon,m} (\log(q^{m-1}))^{1-\varepsilon}, \quad \text{for any } \varepsilon > 0.$$

**Remark 1.25.** The assumption that $m$ is odd is technical, to get the integral monodromy in Theorem 5.12. It is anyway mild, since if $m$ is even, one may as well look at shorter intervals of odd size $m-1$.

**Remark 1.26.** Again, if one wants explicit dependency of $m$, $p$ in the base of $t$ in Theorem 1.20, at the price of a weaker error, one may replace the latter by

$$O_{p,m}\left(\frac{(C(m+1)^{m+3})^t \log\log q}{\log q}\right)$$

with $C$ absolute.

**1D.** ***Outline of the strategy, previous works, and organization of the paper.*** The existence and properties of the limiting distribution under linear independence hypotheses (Theorems 1.14, 1.15, 1.21 and 1.23) follow the methods developed in [Rubinstein and Sarnak 1994; Cha 2008; Martin and Ng 2017]. The continuity statement in Corollary 1.17, under weaker results than full linear independence, is obtained through an idea of Devin [Devin 2019; Devin and Meng 2018].

The main results are then Theorems 1.1, 1.11 and 1.20 on generic linear independence. Combining his large sieve for Frobenius over finite fields [Kowalski 2006; 2008a] with a method of Girstmair [1982; 1999], Kowalski [2008b] proved that a linear independence condition holds generically in some families of $L$-functions of curves over finite fields. This was recently extended by Cha, Fiorilli and Jouve [Cha et al. 2017] to certain families of elliptic curves over function fields, where the underlying symmetry is orthogonal instead of being symplectic.

We use similar ideas to prove Theorems 1.1 and 1.11, with the families of curves replaced by families of exponential sums or characters. More precisely, by work of Deligne [SGA $4\frac{1}{2}$ 1977] and Katz [2017], there are families of $\ell$-adic sheaves on $\mathbb{G}_m$ (resp. on a variety parametrizing primitive characters $\Xi$ or $\chi$ as above) such that the (reversed) characteristic polynomial of the Frobenius acting on a stalk yields the roots (resp. $L$-function) of the corresponding exponential sums (resp. characters).

Unlike in [Kowalski 2008b; Cha et al. 2017], these are not sheaves of $\mathbb{Z}_\ell$-modules, but of $\mathcal{O}_\lambda$-modules, for $\lambda$ a valuation on the ring of integers $\mathcal{O}$ of a number field. In [Kowalski 2008b; Cha et al. 2017] the monodromy structure is symplectic or orthogonal (the latter being the source of complications handled by Jouve); here, it is either special linear, symplectic or projective general linear.

Another difficulty arises in bounding sums of Betti numbers appearing in the large sieve for Frobenius, because certain sheaves are not defined on curves nor have tame ramification, as assumed by Kowalski and Cha, Fiorilli and Jouve. This yields Theorem 5.14, and answers in this case a question of Kowalski [2006, Remark 4.8].

To apply this variant of the large sieve for Frobenius, we also need information on integral monodromy groups of the sheaves, whereas only information about the monodromy groups over $\mathbb{C}$ (i.e., after taking a Zariski closure) is a priori available from Katz's work [1988; 1990; 2013b; 2017]. This is overcome using deep results of Larsen and Pink through ideas of Katz (or more precise results in the case of Kloosterman sums). Unlike in [Cha et al. 2017], strong approximation for arithmetic groups cannot be used.

**Remark 1.27** (Frobenius tori). As explained in [Kowalski 2008b, Section 7], another way to get generic linear independence results is by applying an effective version of Chebotarev's density theorem with Serre's theory of Frobenius tori. However, as explained in [Kowalski 2008b, p. 54], controlling the uniformity with respect to the size of the subsets/tuples considered (crucial for the questions we consider) is more subtle.

**Remark 1.28** (prime polynomials in arithmetic progressions). Keating and Rudnick [2014] also studied the variance of prime polynomials in arithmetic progressions, and obtained as well an asymptotic expression (see [Keating and Rudnick 2014, Theorem 2.2]). In one of the ranges, this uses another

equidistribution result of Katz [2013a]. The latter is more complicated, relying on the ideas developed in [Katz 2012a], because the family involved is not parametrized by an algebraic variety. While results similar to those of Section 1C could probably be obtained (see also [Cha 2008]), we leave that to future work for this reason.

In Sections 2, 3 and 4, respectively for Kloosterman/Birch sums, Gaussian prime polynomials, and prime polynomials in short intervals, we:

(1) Give the cohomological interpretations due to Katz, which gives rise to the eigenvalues from (4), (10) and (12) respectively.

(2) For Gaussian prime polynomials and prime polynomials in short intervals:
   (a) Show the existence of the limiting distributions (Theorems 1.14 and 1.21).
   (b) Prove the additional properties of the distributions under Hypotheses 1.10 and 1.19 (Theorems 1.15 and 1.23).

(3) Prove Corollaries 1.4 and 1.8, 1.7 and Corollaries 1.17, 1.24, from the generic linear independence Theorems 1.1, 1.11 and 1.20 respectively.

Finally, Sections 5, 6 and 7 are dedicated to proving these generic linear independence statements.

**1E.** ***Notations.*** For a prime $p \geq 7$ and a field $E$ with ring of integers $\mathcal{O}$, we let $\mathrm{Spec}_1(\mathcal{O})$ (resp. $\mathrm{Spec}_p(\mathcal{O})$) be the set of all nonzero prime ideals (equivalently, valuations on $\mathcal{O}$) having degree 1 (resp. not lying above $p$), and $\mathrm{Spec}_{1,p}(\mathcal{O}) = \mathrm{Spec}_1(\mathcal{O}) \cap \mathrm{Spec}_p(\mathcal{O})$. If $\lambda \in \mathrm{Spec}_{1,p}(\mathcal{O})$, we denote by $E_\lambda, \mathcal{O}_\lambda$ the completions, and $\mathbb{F}_\lambda \cong \mathcal{O}/\lambda$ the residue field. Note that $\mathbb{F}_\lambda \cong \mathbb{F}_\ell$, where $\ell$ is the prime above which $\lambda$ lies.

## 2. Kloosterman sums and Birch sums

### 2A. *Cohomological interpretation.*

**Theorem 2.1** (Deligne, Katz). *Let $E = \mathbb{Q}(\zeta_{4p})$, with ring of integers $\mathcal{O}$. For every $\lambda \in \Lambda := \mathrm{Spec}_p(\mathcal{O})$, there exists*

(1) *for every integer $r \geq 2$, a lisse sheaf $\mathcal{K}l_{r,\lambda}$ on $\mathbb{G}_{m,\mathbb{F}_p}$ of free $\mathcal{O}_\lambda$-modules, of rank $r$, pure of weight 0, such that for every finite field $\mathbb{F}_q$ of characteristic $p$ and $x \in \mathbb{F}_q^\times$,*

$$\mathrm{tr}(\mathrm{Frob}_{\mathbb{F}_q} \mid (\mathcal{K}l_{r,\lambda})_x) = \mathrm{Kl}_{r,q}(x),$$

*the normalized hyper-Kloosterman sum of rank $r$ defined in* (1)*. Moreover, the family $(\mathcal{K}l_{r,\lambda})_{\lambda\in\Lambda}$ forms a compatible system.*[3]

[3]We recall that this means that for every $\lambda \in \Lambda$, every finite field $\mathbb{F}_q$ of characteristic $p$ and every $x \in \mathbb{F}_q^\times$, the reverse characteristic polynomial $\det(1 - T\,\mathrm{Frob}_{\mathbb{F}_q} \mid (\mathcal{K}l_{r,\lambda})_x) \in \mathcal{O}_\lambda[T]$ has coefficients in $E$ that moreover do not depend on $\lambda$; see [Katz 2001, Section II].

(2) *a lisse sheaf* $\mathcal{B}i_\lambda$ *on* $\mathbb{G}_{m,\mathbb{F}_p}$ *of free* $\mathcal{O}_\lambda$*-modules, of rank* 2*, pure of weight* 0*, such that for every field* $\mathbb{F}_q$ *of characteristic* $p$ *and* $x \in \mathbb{F}_q^\times$,

$$\mathrm{tr}(\mathrm{Frob}_{\mathbb{F}_q} \mid (\mathcal{B}i_\lambda)_x) = \mathrm{Bi}_q(x),$$

*the normalized Birch sum defined in* (2)*. Moreover, the family* $(\mathcal{B}i_\lambda)_{\lambda\in\Lambda}$ *forms a compatible system.*

*Proof.* (1) This is [Katz 1988, Theorem 4.1.1/Section 8.9]. To normalize by a Tate twist, we enlarge the ring of definition to $\mathbb{Z}[\zeta_{4p}]$, which is enough since $\sqrt{p} \in \mathbb{Z}[\zeta_{4p}]$ by the evaluation of quadratic Gauss sums (see [Katz 1988, 11.0]).

(2) This is contained in [Katz 1990, 7.12] (see also [Katz 1987, Part 3]), along with [Katz 1988] for the definition over $\mathcal{O}_\lambda$ of the $\ell$-adic Fourier transform. □

The roots of the characteristic polynomial of $\mathrm{Frob}_{\mathbb{F}_q}$ acting on the stalks at $x \in \mathbb{F}_q^\times$ of any of the sheaves in the system $(\mathcal{K}l_{r,\lambda})_{\lambda\in\Lambda}$, resp. $(\mathcal{B}i_\lambda)_{\lambda\in\Lambda}$, are then the $e(\theta_{i,f,q}(x)) \in \mathbb{C}$ $(1 \le i \le r)$ giving (4), when $f = \mathrm{Kl}_{r,q}$, resp. $f = \mathrm{Bi}_q$ $(r = 2)$.

We now prove the three corollaries of Theorem 1.1 (generic linear independence of the roots) stated in Section 1A.

**2B.** ***Joint uniform distribution: Corollaries 1.4 and 1.7.*** Let $a, b \in \mathbb{F}_q^\times$ be such that the conclusion of Theorem 1.1 holds. By the Kronecker–Weyl equidistribution theorem (see, e.g., [Devin 2019, Section 4.1] or [Martin and Ng 2017, Appendix B]), the random vector

$$\big(n\theta_{i,f,q}(a),\ n\theta_{j,f,q}(b) : 1 \le i, j \le r\big)_{n\le N}$$

equidistributes in $[0, 1]^{2r}$ as $N \to \infty$. It follows at once by (4) that $X_{a,b}$ converges in law to a pair of independent random variables distributed like (8) as $N \to \infty$.

Finally, the equivalence of the distribution of (8) and traces of large enough powers of matrices in $G_r(\mathbb{C})$ is the content of [Rains 1997, Theorem 2.1].

Corollary 1.7 is then an immediate consequence, by applying the portmanteau theorem to the random variable $\mathrm{tr}(g_1) - \mathrm{tr}(g_2)$ (or its real/imaginary parts), which is symmetric around its mean 0. □

**2C.** ***Lower bounds: Corollary 1.8.*** We follow the method of Bombieri and Katz [2010, Sections 3–4], based on the subspace theorem from [Evertse 1984; van der Poorten and Schlickewei 1991] and the Baker–Wüstholz theorem [1993].

Let $a, b \in \mathbb{F}_q^\times$ be such that the conclusion of Theorem 1.1 holds. By (4), we have

$$F(n) := f_{q^n}(a) - f_{q^n}(b) = \sum_{i=1}^{r} \big(e(n\theta_{i,f,q}(a)) - e(n\theta_{i,f,q}(b))\big).$$

The Skolem–Mahler–Lech theorem (see [Bombieri and Katz 2010, Theorem 2.1(i)]) shows that if none of

$$e\left(\frac{\theta_{i,f,q}(x)}{\theta_{j,f,q}(x)}\right) \quad (x \in \{a, b\},\ 1 \le i < j \le r), \qquad e\left(\frac{\theta_{i,f,q}(a)}{\theta_{j,f,q}(b)}\right) \quad (1 \le i, j \le r)$$

are roots of unity, which holds by linear independence, then there are only finitely many $n$ (with $a, b, r, q$ fixed) such that $F(n) = 0$.

The subspace theorem [Evertse 1984; van der Poorten and Schlickewei 1991] (see [Bombieri and Katz 2010, Theorem 3.1]) shows that, after multiplying by $q^{n(r-1)/2}$ (i.e., de-normalizing), for every $n \geq 1$ large enough (with respect to the roots $\theta_{i,f,q}$, i.e., with respect to $a, b, r, q, \varepsilon$), either $F(n) = 0$, or $F(n)$ satisfies the lower bound of Corollary 1.8(1). With the above, this proves the first part of the corollary.

For the second part, we assume that $r = 2$. For any integers $k_0, k_1 \in \mathbb{Z}$ and $\theta_0, \theta_1 \in [0, 1]$, we have

$$\begin{aligned}|\cos(2n\pi\theta_0) - \cos(2n\pi\theta_1)| &= 2|\sin(n\pi(\theta_0+\theta_1))\sin(n\pi(\theta_0-\theta_1))| \\ &= 2\prod_{j=0}^{1}|\sin(n\pi\tau_j - k_j\pi)| \qquad (\tau_j = \theta_0 + (-1)^j\theta_1) \\ &\geq 2\prod_{j=0}^{1}\frac{2|n\pi\tau_j - k_j\pi|}{\pi} \\ &= \frac{2}{\pi^2}\prod_{j=0}^{1}|n\log(e(\tau_j)) - k_j\log(-1)|,\end{aligned}$$

where the inequality holds if $k_j$ is chosen to minimize $|n\tau_j - k_j|$.

We can now apply the Baker–Wüstholz theorem [1993, Theorem, p. 20] as in [Bombieri and Katz 2010, Section 4], or its improvement with respect to the numerical constants by Gouillon [2006], giving the first and second expressions in Corollary 1.8(2). As the arguments are essentially the same, we only give the second one. If $1, \theta_0, \theta_1$ are linearly independent, then [Gouillon 2006, Corollary 2.2] shows that this is

$$\geq \frac{2}{\pi^2}\prod_{j=0}^{1}\exp\left(-9400\left(3.317 + \frac{1.888}{d} + 0.946\log d\right)d^4 h_j A_j\right), \tag{13}$$

where $A_j$ is any real number satisfying $\log A_j \geq \max(1, h(e(\tau_j)), |\tau_j|/d, 1/d)$,

$$\begin{aligned}h_j &= \max\left(\log\left(\frac{n}{ed} + \frac{k_j}{dA_1}\right), \frac{1000}{d}, 498 + \frac{284}{d} + 142\log d\right), \\ d &= [\mathbb{Q}(e(\tau_0), e(\tau_1)) : \mathbb{Q}]/2,\end{aligned}$$

for $h_0$ the absolute logarithmic Weil height. We have $h_0(e(\tau_j)) \leq h_0(e(\theta_0)) + h_0(e(\theta_1))$.

Let us now assume that $(\theta_0, \theta_1) = (\theta_{i,f,q}(a), \theta_{i,f,q}(b))$ are moreover angles of exponential sums (4). Then $q^{1/2}e(\pm\theta_j)$ is an algebraic integer, so $h_0(e(\tau_j)) \leq \log q$. Regarding the degree, we have that $1 \leq d \leq (p-1)/2$ as in [Bombieri and Katz 2010, Proof of Corollary 4.3], because Kloosterman/Birch sums are sums of $p$-th roots of unity. Thus, we may take $A_j = \max(q, e^2)$ and

$$h_j \leq \max\left(\log\left(\frac{n}{e} + \frac{2n + \frac{1}{2}}{A_j}\right), 1000, 782 + 142\log\frac{p-1}{2}\right).$$

Then, (13) is

$$\geq \frac{2}{\pi^2}\exp\left(-1175\left(5.205+0.946\log\frac{p-1}{2}\right)(p-1)^4 h\max(\log q,2)\right),$$

where

$$h=\max\left(\log\left(\frac{n}{e}+\frac{2n+\frac{1}{2}}{q}\right),1000,782+142\log\frac{p-1}{2}\right).$$

If $p$ is fixed and $n$ is large enough with respect to it, this gives the expression in Corollary 1.8. This yields the result by Theorem 1.1. The argument is essentially the same to lower bound a single Kloosterman sum with Gouillon's result, with the analogue of (13) having a leading factor of $2/\pi$, no product, and $A_0=\max(q/2,e)$. □

## 3. Angles of Gaussian primes

### 3A. *Definitions and cohomological interpretation.*

**Definition 3.1.** Let $q$ be an odd prime power and $k\geq 2$ be an integer. A *super-even character* $\Xi$ *modulo* $S^k$ *over* $\mathbb{F}_q$ is a character of

$$\mathbb{S}^1_{k,q}\cong R_{k,q}/H_k,\qquad H_k:=(\mathbb{F}_q[S^2]/(S^k))^\times$$

(see (9)). The *Swan conductor* of a nontrivial $\Xi$ is the maximal (odd) integer $d(\Xi)$ such that $\Xi$ is nontrivial on $(1+(S^{d(\Xi)}))/(S^k)\leq R_{k,q}$. The character $\Xi$ is *primitive* if $d(\Xi)=2\kappa-1$, with $\kappa:=\lfloor k/2\rfloor$. The *L-function* of a nontrivial $\Xi$ is

$$L(\Xi,T)=\prod_{\substack{P\text{ prime}\\ \text{monic}\\ P(0)\neq 0}}(1-\Xi(P)T^{\deg P})^{-1}. \tag{14}$$

**Theorem 3.2** (Katz). *Let $\mathbb{F}_q$ be a finite field of odd characteristic $p$, $k\geq 2$ be an even integer,*

$$E=\mathbb{Q}\left(\zeta_{4p^r}:1\leq r\leq 1+\frac{\log k}{\log p}\right)\subset\mathbb{Q}(\zeta_{p^\infty})$$

*with ring of integers $\mathcal{O}$, and let $\lambda\in\Lambda:=\mathrm{Spec}_p(\mathcal{O})$.*

(1) *There exists a unipotent group $\mathbb{W}_{k,\mathrm{odd}}$ over $\mathbb{F}_p$ such that $\mathbb{W}_{k,\mathrm{odd}}(\mathbb{F}_q)=\mathbb{S}^1_{k,q}$ (the group of super-even characters, by duality), as well as an open set $\mathrm{Prim}_{k,\mathrm{odd}}\subset\mathbb{W}_{k,\mathrm{odd}}$ such that $\mathrm{Prim}_{k,\mathrm{odd}}(\mathbb{F}_q)$ is in bijection with primitive super-even characters modulo $S^k$ over $\mathbb{F}_q$.*

(2) *There exists a lisse sheaf $\mathcal{G}_{k,\lambda}$ on $\mathrm{Prim}_{k,\mathrm{odd}}$ of free $\mathcal{O}_\lambda$-modules, of rank $r=2\kappa-2$, pure of weight 1, such that for every $\Xi\in\mathrm{Prim}_{k,\mathrm{odd}}(\mathbb{F}_q)$, we have*

$$\det(1-T\,\mathrm{Frob}_{q,\Xi}\mid\mathcal{G}_{k,\lambda})=\frac{L(\Xi,T)}{1-T},$$

*which is a polynomial of degree $d(\Xi)=r+1$. In particular, the family $(\mathcal{G}_{k,\lambda})_{\lambda\in\Lambda}$ forms a compatible system.*

(3) *The Tate twist* $\mathcal{F}_{k,\lambda} = \mathcal{G}_{k,\lambda}\left(\frac{1}{2}\right)$ *is a lisse sheaf of free* $\mathcal{O}_\lambda$*-modules on* $\mathrm{Prim}_{k,\,\mathrm{odd}}$*, pure of weight zero, of rank* $d(\Xi) - 1$*, with symplectic auto-duality.*

*Proof.* These are the contents of [Katz 2017, Section 2] (see also the constructions in [Katz 2013b, Sections 1–4]). □

In particular, the eigenvalues of $\mathrm{Frob}_{\mathbb{F}_q}$ acting on the stalks of $\mathcal{F}_{k,\lambda}$ at super-even primitive $\Xi$, which are free $\mathcal{O}_\lambda$-modules of rank $2\kappa - 2$, yield the eigenvalues $e(\pm\theta_{\Xi,j}) \in \mathbb{C}$ from (10), such that

$$\begin{aligned} L(\Xi, T) &= (1-T)\prod_{j=1}^{\kappa-1}\big(1-\sqrt{q}e(\theta_{\Xi,j})T\big)\big(1-\sqrt{q}e(-\theta_{\Xi,j})T\big) \\ &= (1-T)\det(1-\sqrt{q}T\Theta_\Xi), \quad \text{with } \Theta_\Xi \in \mathrm{Sp}_{d(\Xi)-1}(\mathbb{C}). \end{aligned}$$

**3B.** ***Existence of the limiting distribution.*** We start with an explicit formula for $X_{k,N}(\boldsymbol{u})$.

**Proposition 3.3.** *For all* $u \in \mathbb{S}^1_{k,q}$ *and* $n \le N$*, we have*

$$X_{k,N}(u)_n = -2\sum_{f=2}^{\kappa}\sum_{j=1}^{f-1}\sum_{\substack{\Xi\in\widehat{\mathbb{S}}^1_{k,q}\\ d(\Xi)=2f-1}} \overline{\Xi}(u)\cos(2\pi n\theta_{\Xi,j}) - \delta_{n\ \mathrm{even}}|\{b\in\mathbb{S}^1_{k,q} : b^2 = u\}| + O\left(\frac{q^{k/2}\tau(n)}{q^{n/6}n} + \frac{kq^k}{q^{n/4}}\right),$$

*with an absolute implied constant. Moreover,* $|\{b \in \mathbb{S}^1_{k,q} : b^2 = u\}| \in \{0, 1\}$ *and in the expression above,* $\overline{\Xi}(u)\cos(2\pi n\theta_{\Xi,j})$ *may be replaced by* $\mathrm{Re}(e(\theta_{\Xi,j})\Xi(u))$.

**Remark 3.4.** Almost all (i.e., a density $1 + O(1/q)$) super even $\Xi \in \widehat{\mathbb{S}}^1_{k,q}$ have conductor $2\kappa - 1$, but since we look at the $N \to \infty$ limit, we cannot restrict the sum in Proposition 3.3 to those characters only as in [Rudnick and Waxman 2019, Proof of Theorem 6.7] (with a $q \to \infty$ limit).

*Proof.* By [Rudnick and Waxman 2019, Lemma 6.4, Section 6.6], we have

$$X_{k,N}(u)_n = -\sum_{\Xi\neq 1}\overline{\Xi}(u)\,\mathrm{tr}\,\Theta_\Xi^n - \frac{R_{k,n}(u)q^\kappa}{q^{n/2}} - \frac{\delta_{u=1}q^\kappa}{q^{n/2}},$$

where, by the prime polynomial theorem [Rosen 2002, Theorem 2.2],

$$R_{k,n}(u) := \sum_{\substack{f\in\mathbb{F}_q[S]\ \text{monic}\\ \text{not prime}\\ \deg(f)=n}} \Lambda(f)\delta_{U(f)\in\mathrm{Sec}(u,k)} = \delta_{n\ \mathrm{even}}\,\frac{n}{2}\sum_{\substack{P\ \text{monic}\\ \text{prime}\\ \deg(P)=n/2}} \delta_{U(P^2)\in\mathrm{Sec}(u,k)} + O\left(\frac{q^{n/3}\tau(n)}{n}\right).$$

By the function field analogue of Dirichlet's theorem on primes in arithmetic progressions [Rosen 2002, Theorem 4.8], if $n$ is even,

$$\begin{aligned}
-\frac{q^\kappa n}{2q^{n/2}} \sum_{\substack{P \text{ monic} \\ \text{prime} \\ \deg(P)=n/2}} \delta_{U(P^2)\in \mathrm{Sec}(u,k)} &= -\frac{q^\kappa n}{2q^{n/2}} \sum_{\substack{a\in R_{k,q} \\ a^2\equiv u \pmod * H_k}} \sum_{\substack{P \text{ monic} \\ \text{prime} \\ \deg(P)=n/2}} \delta_{P\equiv a \pmod * S^k} \\
&= -\frac{q^\kappa n}{2q^{n/2}} \sum_{\substack{a\in R_{k,q} \\ a^2\equiv u \pmod * H_k}} \left(\frac{1}{|R_{k,q}|}\frac{q^{n/2}}{n/2} + O\left(\frac{q^{n/4}k}{n}\right)\right) \\
&= -|\{a\in R_{k,q} : a^2 \equiv u \pmod * H_k\}| \left(\frac{1}{|H_k|} + O\left(\frac{q^\kappa k}{q^{n/4}}\right)\right) \\
&= -|\{b\in \mathbb{S}^1_{k,q} : b^2 = u\}| \left(1 + O\left(\frac{k|R_{k,q}|}{q^{n/4}}\right)\right).
\end{aligned}$$

Note that in odd characteristic, the cardinality $|\mathbb{S}^1_{k,q}| = q^\kappa$ is odd, so the function $(x \in \mathbb{S}^1_{k,q}) \mapsto x^2$ is injective, and $|\{b \in \mathbb{S}^1_{k,q} : b^2 = u\}| \in \{0, 1\}$.

Hence,

$$X_{k,N}(u)_n = -\sum_{\Xi\neq 1} \overline{\Xi}(u) \operatorname{tr} \Theta^n_\Xi - \frac{\delta_{u=1}q^\kappa}{q^{n/2}} + O\left(\frac{q^\kappa \tau(n)}{q^{n/6}n}\right) - \delta_{n \text{ even}} |\{b\in \mathbb{S}^1_{k,q} : b^2 = u\}| \left(1 + O\left(\frac{kq^k}{q^{n/4}}\right)\right),$$

which gives the result after splitting the sum over characters $\Xi$ depending on the conductors $d(\Xi)$, which are odd integers. The last assertion follows from the invariance of the sum under $\Xi \mapsto \overline{\Xi}$. □

*Proof of Theorem 1.14.* The existence of the limiting distribution goes almost exactly as in [Cha 2008, Lemma 3.1, Theorem 3.2] (based on [Rubinstein and Sarnak 1994]). Let $\tilde{X}_{k,N}(\boldsymbol{u})$ be the random variable on $[1, N]$ defined by the right-hand side of the expression in Proposition 3.3, but without the error term. Let moreover

$$V := \big\{(\Xi, j) : \Xi \in \widehat{\mathbb{S}}^1_{k,q},\ \Xi \neq 1,\ 1 \le j \le d'(\Xi)\big\}. \tag{15}$$

There exists an explicit continuous function $g_{k,\boldsymbol{u}} : (\mathbb{R}/\mathbb{Z})^V \to \mathbb{R}^R$ such that

$$\tilde{X}_{k,N}(\boldsymbol{u}) = \big(g_{k,\boldsymbol{u}}(n\theta_{\Xi,j} : (\Xi, j) \in V)\big)_{n\le N.}$$

Note that $g_{k,\boldsymbol{u}}$ is bounded is (when $k, q$ are fixed): each component is bounded by $2\kappa q^\kappa$.

By the Kronecker–Weyl equidistribution theorem, $\big(n\theta_{\Xi,j} : (\Xi, j) \in V\big)_{n\le N}$ converges in law (as $N \to \infty$) to a random vector equidistributed in the closure $\overline{\Gamma}$ of the torus

$$\Gamma = \{n(\theta_{\Xi,j})_{(\Xi,j)\in V} : n \in \mathbb{Z}\} \subset (\mathbb{R}/\mathbb{Z})^V. \tag{16}$$

It then follows from Helly's selection theorem [Billingsley 1986, Theorems 25.9–10] that $X_{k,N}(\boldsymbol{u})$ converges in law to a random vector $X_k(\boldsymbol{u})$ which corresponds to a measure $\mu_{k,\boldsymbol{u}}$ satisfying

$$\int_{\mathbb{R}^R} f(\boldsymbol{x})\, d\mu_{k,\boldsymbol{u}}(\boldsymbol{x}) = \int_{\overline{\Gamma}} (f \circ g_{k,\boldsymbol{u}})(\boldsymbol{x})\, d\boldsymbol{x} \tag{17}$$

for every bounded continuous $f : \mathbb{R}^R \to \mathbb{R}$. The limiting measure $\mu_{k,\boldsymbol{u}}$ is compactly supported from the boundedness of $g_{k,\boldsymbol{u}}$ ($k, q$ fixed).

In particular, there is convergence of the moments, which allows us to compute the expected value by noting that

$$\left| \frac{1}{N} \sum_{f=2}^{\kappa} \sum_{j=1}^{f-1} \sum_{\substack{\Xi \in \widehat{\mathbb{S}}^1_{k,q} \\ d(\Xi)=2f-1}} \overline{\Xi}(u) \sum_{n=1}^{N} \cos(2\pi n \theta_{\Xi,j}) \right| \ll \frac{\kappa q^{\kappa}}{N} \xrightarrow{N\to\infty} 0. \qquad \square$$

**3C.** ***Properties of the limiting distribution under (generic) linear independence.*** For the next properties, we continue to use the methods of Rubinstein and Sarnak [1994] and others, in particular by studying characteristic functions.

**Lemma 3.5** (Fourier transform). *For $u_1, \dots, u_R \in \mathbb{S}^1_{k,q}$ distinct, let $\mu_{k,\boldsymbol{u}}$ be the measure associated with the $R$-dimensional random vector $X_k(\boldsymbol{u})$. Its Fourier transform*

$$\hat{\mu}_{k,\boldsymbol{u}}(\boldsymbol{t}) := \int_{\mathbb{R}^R} e^{-i\boldsymbol{t}\cdot\boldsymbol{x}}\, d\mu_{k,\boldsymbol{u}}(\boldsymbol{x}) \qquad (\boldsymbol{t} \in \mathbb{R}^R)$$

*is given by*

$$\exp(i\boldsymbol{t}\cdot\boldsymbol{b}_k(\boldsymbol{u})) \int_{\overline{\Gamma}} \prod_{f=1}^{\kappa} \prod_{j=1}^{f-1} \prod_{\substack{\Xi \in \widehat{\mathbb{S}}^1_{k,q} \\ d(\Xi)=2f-1}} \exp\big(2i \operatorname{Re}(e(x_j)\boldsymbol{t}\cdot\Xi(\boldsymbol{u}))\big)\, d\boldsymbol{x},$$

*where $\Gamma$ is the torus* (16) *and $\boldsymbol{b}_k(\boldsymbol{u}) := (|\{b \in \mathbb{S}^1_{k,q} : b^2 = u_r\}|/2)_{1\le r\le R}$, $\Xi(\boldsymbol{u}) := (\Xi(u_r))_{1\le r\le R}$. If Hypothesis 1.10 holds, then*

$$\hat{\mu}_{k,\boldsymbol{u}}(\boldsymbol{t}) = \exp(i\boldsymbol{t}\cdot\boldsymbol{b}_k(\boldsymbol{u})) \prod_{f=2}^{\kappa} \prod_{j=1}^{f-1} \prod_{\substack{\Xi \in \widehat{\mathbb{S}}^1_{k,q} \\ d(\Xi)=2f-1}} J_0(2|\boldsymbol{t}\cdot\Xi(\boldsymbol{u})|), \tag{18}$$

*where $J_0(z) = \frac{1}{\pi}\int_0^{\pi} \cos(z\sin t)\, dt$ is the 0-th Bessel function of the first kind.*

*Proof.* The first statement is a direct consequence of Proposition 3.3 and (17). For (18), under Hypothesis 1.10 the torus $\overline{\Gamma}$ is maximal and the integral splits as a product of integrals of the form

$$\int_{\mathbb{R}/\mathbb{Z}} \exp\big(2i \operatorname{Re}(e(x_j)\boldsymbol{t}\cdot\Xi(\boldsymbol{u}))\big)\, dx_j = J_0(2|\boldsymbol{t}\cdot\Xi(\boldsymbol{u})|)$$

by [Martin and Ng 2017, Lemma C.1]. $\square$

We now prove Theorem 1.15 about properties of the limiting distribution under Hypothesis 1.10.

*Proof of Theorem 1.15.* To show that $X_k(\boldsymbol{u})$ is absolutely continuous, it is enough to show that $\int_{\mathbb{R}^R} |\hat{\mu}_{k,\boldsymbol{u}}(\boldsymbol{t})|\, d\boldsymbol{t} < \infty$; see [Martin and Ng 2017, Lemma A.8(b)]. To do so, we partly follow the method of [Martin and Ng 2017, Section 4]. Since we assume Hypothesis 1.10, we may use (18) from Lemma 3.5:

$$|\hat{\mu}_{k,\boldsymbol{u}}(\boldsymbol{t})| \le \prod_{f=2}^{\kappa} \prod_{j=1}^{f-1} \prod_{\substack{\Xi \in \widehat{\mathbb{S}}^1_{k,q} \\ d(\Xi)=2f-1}} |\boldsymbol{t}\cdot\Xi(\boldsymbol{u})|^{-1/2} \le \Big[ \prod_{\Xi\in S} |\boldsymbol{t}\cdot\Xi(\boldsymbol{u})|^2 \Big]^{-\frac{\kappa-1}{4}}$$

$$\le \Big[ \frac{1}{|S_1(\boldsymbol{t})|} \sum_{\Xi\in S_1(\boldsymbol{t})} |\boldsymbol{t}\cdot\Xi(\boldsymbol{u})|^2 \Big]^{-\frac{\kappa-1}{4}}$$

where

$$S_1(\boldsymbol{t}) := \{\Xi \in \widehat{\mathbb{S}}^1_{2\kappa,q} \text{ primitive} : |\boldsymbol{t}\cdot\Xi(\boldsymbol{u})| > 1\} \subset S := \{\Xi \in \widehat{\mathbb{S}}^1_{2\kappa,q} \text{ primitive}\},$$

since $|J_0(z)| \le \min(1, \sqrt{2/(\pi|z|)})$ for all $z \in \mathbb{R}$ (see [Martin and Ng 2017, Lemma C.2]). If $\boldsymbol{t} \in \boldsymbol{T} := \{\boldsymbol{t} \in \mathbb{R}^R : |S_1(\boldsymbol{t})| \ge 1\}$, we get

$$\frac{1}{|S_1(\boldsymbol{t})|} \sum_{\Xi\in S_1(\boldsymbol{t})} |t\cdot\Xi(\boldsymbol{u})|^2 \ge \frac{1}{|S|} \sum_{\Xi\in S} |t\cdot\Xi(\boldsymbol{u})|^2 = \sum_{r,r'=1}^{R} t_r \overline{t_{r'}} \frac{1}{|S|} \sum_{\Xi\in S} \Xi(u_r)\overline{\Xi}(u_{r'}).$$

By the orthogonality relations and Möbius inversion,

$$\frac{1}{|S|} \sum_{\Xi\in S} \Xi(u_r)\overline{\Xi}(u_{r'}) = \frac{1}{|S|} \sum_{f=2}^{\kappa} \mu(S^{2(\kappa-f)}) \sum_{\Xi\in\widehat{\mathbb{S}}^1_{2f,q}} \Xi(u_r)\overline{\Xi}(u_{r'}) = \frac{q^\kappa \delta_{u_r=u_{r'}}}{|S|} = \frac{\delta_{u_r=u_{r'}}}{1-1/q}.$$

Since the $u_i$ are distinct, it follows that

$$\frac{1}{|S_1(\boldsymbol{t})|} \sum_{\Xi\in S_1(\boldsymbol{t})} |\boldsymbol{t}\cdot\Xi(\boldsymbol{u})|^2 \ge \|\boldsymbol{t}\|^2 \qquad \text{if } |S_1(\boldsymbol{t})| \ge 1.$$

Therefore, if $\boldsymbol{t} \in \boldsymbol{T}$, then $|\hat{\mu}_{k,\boldsymbol{u}}(\boldsymbol{t})| \le \|\boldsymbol{t}\|^{-(\kappa-1)/2}$. On the other hand, if $\boldsymbol{t} \notin \boldsymbol{T}$, the same argument shows that

$$1 \ge \frac{1}{|S|} \sum_{\Xi\in S} |\boldsymbol{t}\cdot\Xi(\boldsymbol{u})| \ge \|\boldsymbol{t}\|^2,$$

i.e., $\mathbb{R}^R \backslash \boldsymbol{T}$ is bounded. It also contains a neighborhood of $\mathbf{0}$ since it contains the finite intersection $\bigcap_{\Xi\in S}\{\boldsymbol{t} \in \mathbb{R}^R : |\boldsymbol{t}\cdot\Xi(\boldsymbol{u})| < 1\}$ of open sets containing $\mathbf{0}$.

Thus, there exists $\varepsilon > 0$ such that

$$\int_{\mathbb{R}^R} |\hat{\mu}_{k,\boldsymbol{u}}(\boldsymbol{t})|\, d\boldsymbol{t} \ll \int_{\|\boldsymbol{t}\|\le 1} |\hat{\mu}_{k,\boldsymbol{u}}(\boldsymbol{t})|\, d\boldsymbol{t} + \int_{\mathbb{R}^R \backslash B_\varepsilon(\mathbf{0})} \|\boldsymbol{t}\|^{-(\kappa-1)/2}\, d\boldsymbol{t},$$

and the second integral converges when $\kappa - 1 > 2R$; see [Martin and Ng 2017, p. 22]. This concludes the proof of (1).

Concerning (2), the symmetry/exchangeability follow from the expression (18) for $\hat{\mu}_{k,\boldsymbol{u}}$.

The last statements of the theorem follow from the previous ones: since $\mu_{k,\boldsymbol{u}}$ is absolutely continuous, $A = \{\boldsymbol{x} \in \mathbb{R}^R : x_1 < \cdots < x_R\}$ is a continuity set, so that by the portmanteau theorem,

$$\lim_{N\to\infty} \mathbb{P}\big(X_{k,N}(u_1) < \cdots < X_{k,N}(u_R)\big) = \mu_{k,\boldsymbol{u}}(A). \qquad \square$$

Finally, we prove Corollary 1.17 (unconditional properties of the limiting distribution) assuming Theorem 1.11 on generic linear independence.

*Proof of Corollary 1.17.* (1) It suffices to show it when $R = 1$, i.e., that the random variable $X_k(u)$ is continuous for every $u \in \mathbb{S}^1_{k,q}$. We follow the argument in [Devin 2019, Proof of Theorem 2.2]; see also [Devin and Meng 2018, Proposition 2.1]. By Wiener's lemma, it suffices to show that

$$\lim_{S\to\infty} \frac{1}{S} \int_{-S}^{S} |\hat{\mu}_{k,u}(t)|^2 \, dt = 0. \tag{19}$$

By Lemma 3.5, $|\hat{\mu}_{k,u}(t)| \le |\int_{\overline{\Gamma}} \exp(it\phi(\boldsymbol{x}))\, d\boldsymbol{x}|$, where

$$\phi(\boldsymbol{x}) := 2 \sum_{f=1}^{\kappa} \sum_{j=1}^{f-1} \sum_{\substack{\Xi \in \widehat{\mathbb{S}}^1_{k,q} \\ d(\Xi)=2f-1}} \cos(2\pi x_j)\, \Xi(u).$$

By Theorem 1.11, there exists $\Xi \in \widehat{\mathbb{S}}^1_{k,q}$ and $1 \le j \le d'(\Xi)$ such that $\theta_{\Xi,j} \notin \mathbb{Q}$. It follows that the function $\phi : \overline{\Gamma} \to \mathbb{R}$ is analytic and nonconstant, since $\Xi(u) \ne 0$ (being a root of unity). Thus, the scaling principle [Stein 1993, VIII.2, Proposition 5] shows that $|\hat{\mu}_{k,u}(t)| \ll |t|^{-\alpha}$ for some constant $\alpha > 0$, where $\alpha$ and the implied constant can depend on all parameters but $t$. Thus, (19) holds, using the trivial bound $|\hat{\mu}_{k,u}(t)| \le 1$ around 0.

(2) This is a consequence of the proof of Theorem 1.14: $\overline{\Gamma}$ is a subtorus of $(\mathbb{R}/\mathbb{Z})^V$, with $V$ as in (15), and if the set of the $\theta_{\Xi,j}$ ($(\Xi, j) \in V$) contains at least $t$ linearly independent elements, then $\dim \overline{\Gamma} \ge t$. By Theorem 1.11, the latter holds whenever $t = o(\log |\mathbb{S}^1_{k,q}|)$. $\square$

## 4. Prime polynomials in short intervals

### 4A. *Definitions and cohomological interpretation.*

**Definition 4.1.** Let $Q \in \mathbb{F}_q[T]$ be nonconstant.

- A Dirichlet character $\chi$ modulo $Q$ is a character of $(\mathbb{F}_q[T]/(Q))^\times$.
- The character $\chi$ is *even* if it is trivial on $\mathbb{F}_q^\times$.
- It is *primitive* if it is not induced from a character modulo a proper divisor $Q' \mid Q$ through the natural map $(\mathbb{F}_q[T]/(Q))^\times \to (\mathbb{F}_q[T]/(Q'))^\times$. The *conductor* of $\chi$ is the monic divisor $Q' \mid Q$ of smallest degree such that $\chi$ is primitive modulo $Q'$.
- As usual, we may extend $\chi$ as $\chi : \mathbb{F}_q[T] \to \mathbb{C}$ by defining $\chi(f) = \chi(f \pmod{*Q})$ if $(f, Q) = 1$, $\chi(f) = 0$ otherwise.

- The number of Dirichlet characters modulo $Q$ is denoted by $\varphi(Q)$. The number of even (resp. primitive, even primitive) such characters is $\varphi^{\mathrm{ev}}(Q) = \varphi(Q)/(q-1)$ (resp. $\varphi_{\mathrm{prim}}(Q)$, $\varphi^{\mathrm{ev}}_{\mathrm{prim}}(Q)$).
- The *$L$-function* of $\chi$ is
$$L(\chi, T) = \prod_{\substack{P \text{ prime} \\ \text{monic} \\ P \nmid Q}} (1 - \chi(P) T^{\deg P})^{-1}.$$

We recall that if $\deg(Q) \geq 2$ and $\chi \neq 1$, then $L(\chi, T)$ is a polynomial (rather than a formal power series) of degree $\deg(Q) - 1$; see [Rosen 2002, Proposition 4.3 and p. 130].

If $\chi$ is even, then $L(\chi, T)$ has a "trivial" zero at $T = 1$. As in [Keating and Rudnick 2014, (3.34)], we define $\lambda_\chi = \delta_{\chi \text{ even}}$, which allows to factor

$$L(\chi, T) = (1 - \lambda_\chi T) L^*(\chi, T), \quad L^*(\chi, T) \in \mathbb{F}_q[T].$$

If $\chi$ is primitive, Weil's work on the Riemann hypothesis over finite fields (see [Rosen 2002, Chapters 4, 5]) shows that

$$L^*(\chi, T) = \det(1 - \sqrt{q} T \Theta_\chi), \quad \Theta_\chi \in U_{\deg(Q)-1-\lambda_\chi}(\mathbb{C}), \tag{20}$$

and we let

$$e(\theta_{\chi,j}), \quad (1 \leq j \leq \deg(Q) - 1 - \lambda_\chi), \quad \theta_{\chi,j} \in [0, 1],$$

be the eigenvalues of $\Theta_\chi^{-1}$. This is also reflected in the following result:

**Theorem 4.2** (Katz). *Let $\mathbb{F}_q$ be a finite field of odd characteristic $p$, $m \geq 2$ be an integer,*

$$E = \mathbb{Q}\left(\zeta_{m-2}, \zeta_{4p^r} : 1 \leq r \leq 1 + \frac{\log m}{\log p}\right) \subset \mathbb{Q}(\zeta_{p^\infty}, \zeta_n)$$

*with ring of integers $\mathcal{O}$, and let $\lambda \in \Lambda := \mathrm{Spec}_p(\mathcal{O})$.*

(1) *There exists a unipotent group $\mathbb{W}_m$ over $\mathbb{F}_p$ such that $\mathbb{W}_m(\mathbb{F}_q)$ is the group of even characters modulo $T^m \in \mathbb{F}_q[T]$, as well as an open set $\mathrm{Prim}_m \subset \mathbb{W}_m$ such that $\mathrm{Prim}_m(\mathbb{F}_q)$ is the set of primitive even characters modulo $T^m$.*

(2) *There exists a lisse sheaf $\mathcal{G}_{m,\lambda}$ on $\mathrm{Prim}_m$ of free $\mathcal{O}_\lambda$-modules, of rank $m-2$, pure of weight 1, such that for every $\chi \in \mathrm{Prim}_m(\mathbb{F}_q)$,*

$$\det(1 - T \operatorname{Frob}_{q,\chi} \mid \mathcal{G}_{m,\lambda}) = L^*(\chi, T),$$

*which is a polynomial of degree $m-2$. In particular, the family $(\mathcal{G}_{m,\lambda})_{\lambda \in \Lambda}$ forms a compatible system.*

(3) *The Tate twist $\mathcal{F}_{m,\lambda} = \mathcal{G}_{m,\lambda}\left(\frac{1}{2}\right)$ is a lisse sheaf of free $\mathcal{O}_\lambda$-modules on $\mathrm{Prim}_m$, pure of weight zero, of rank $m-2$.*

In other words, the eigenvalues of $\sqrt{q}\Theta_\chi$ (the zeros of $L^*(\chi, T)$) are the eigenvalues of $\mathrm{Frob}_{\mathbb{F}_q}$ acting on the stalk of $\mathcal{G}_{m,\lambda}$ at $\chi$.

*Proof.* This is essentially the contents of [Katz 2013b, Sections 1–4]. The addition of $\zeta_{m-2}$ is not necessary at this point, but will be useful in Theorem 5.12. □

**4B.** ***Existence of the limiting distribution.*** We start with an explicit formula for $X_{m,N}(\boldsymbol{B})$, and proceed as in Section 3B.

**Proposition 4.3.** *Under the notations of Section 1C, we have, for $B \in \mathbb{F}_q[T]$ monic of degree $m-1$,*

$$X_{m,N}(B)_n = -\sum_{f=3}^{m} \sum_{\substack{\chi \pmod * T^m \text{ even} \\ \mathrm{cond}(\chi)=T^f}} \sum_{j=1}^{f-2} \overline{\chi}(B^*) e(\theta_{\chi,j}) + \frac{1}{q^{n/2}},$$

*where $B^* \in \mathbb{F}_q[T]$ is the reflected polynomial defined by $B^*(T) = T^{\deg B} B(1/T)$.*

*Proof.* By [Keating and Rudnick 2014, (4.22)],

$$X_{m,N}(B)_n = \frac{1}{q^{n/2}} \sum_{\substack{\chi \pmod * T^m \\ \text{even}}} \overline{\chi}(B^*)\psi(n,\chi), \qquad \psi(n,\chi) := \sum_{\substack{f \in \mathbb{F}_q[T] \\ \deg(f)=n}} \Lambda(f)\chi(f) = -q^{n/2}\,\mathrm{tr}(\Theta_\chi^n) - 1,$$

where the last equality is the explicit formula for $\psi$ (see [Keating and Rudnick 2014, (3.38)]), obtained by taking the logarithmic derivative on both sides of (20). Thus,

$$X_{m,N}(B)_n = \frac{1}{q^{n/2}} \sum_{\substack{\chi \pmod * T^m \\ \text{even}}} \overline{\chi}(B^*)\,\mathrm{tr}(\Theta_\chi^n) - \frac{1}{q^{n/2}} \sum_{\substack{\chi \pmod * T^m \\ \text{even}}} \overline{\chi}(B^*).$$

The result follows after splitting the first sum according to the conductor of $\chi$ and applying the orthogonality relations in $(\mathbb{F}_q[T]/(T^m))^\times/\mathbb{F}_q^\times$ to the second sum. □

Then, the proof of Theorem 1.21 is exactly like the proof of Theorem 1.14 (see Section 3B). As in Proposition 3.3, one may replace the $e(\theta_{\chi,j})$ in Proposition 4.3 by $\cos(2\pi\theta_{\chi,j})$ since $X_{m,N}(B)_n \in \mathbb{R}$.

**4C.** ***Properties of the limiting distribution under (generic) linear independence.*** Again, the proofs of Theorem 1.23 and Corollary 1.24 are exactly like the proofs of Theorem 1.15 and Corollary 1.17 respectively, in Section 3C.

## 5. An extension of the large sieve for Frobenius

In the next two sections, we set up the tools to prove the main Theorems 1.1, 1.11 and 1.20 on generic linear independence. As outlined in Section 1D, the strategy follows that of previous works and is the following:

(1) Obtain information about integral monodromy groups of reductions of sheaves of $\mathcal{O}_\lambda$-modules from Theorem 2.1 and 3.2, for a set of ideals/valuations $\lambda \in \mathrm{Spec}_{1,p}(\mathcal{O})$ of positive density.

(2) Use a variant of the large sieve for Frobenius to show that for all such $\lambda$, the (splitting) fields generated by the roots ($\alpha_{i,f,p}(x)$, $e(\theta_{\Xi,j})$ or $e(\theta_{\chi,j})$) are maximal for almost all tuples of arguments $x$ (resp. $\Xi$, $\chi$) for exponential sums (resp. (super-)even characters).

(3) Apply Girstmair's work to show that (2) implies the desired linear independence.

The first two points and the variant of the large sieve for Frobenius are implemented in this section, and the third point in Section 6.

**Remark 5.1.** Note that [Kowalski 2008b; Cha et al. 2017] dealt with symplectic and orthogonal monodromy types. Here, we need to consider special linear and symplectic ones, which will correspond to splitting fields with Galois groups $\mathfrak{S}_n$ (the full symmetric group), or $W_{2n} \leq \mathfrak{S}_{2n}$, the subgroup with order $2^n n!$ of permutations of $n$ pairs (the Coxeter group $B_n$).

**Remark 5.2.** We consider ideals of degree 1 so that $\mathbb{F}_\lambda = \mathbb{F}_\ell$ and considerations on the sheaves mod $\lambda$ can be reduced as much as possible to existing arguments, for the large sieve or computations of integral monodromy groups. This is actually not a restriction because $\mathrm{Spec}_{1,p}(\mathcal{O})$ has natural density 1 in $\mathrm{Spec}(\mathcal{O})$ [Narkiewicz 2004, Corollary 2, p. 345, Proposition 7.17].

**Remark 5.3.** Since we considered Tate-twisted/normalized sheaves of $\mathcal{O}_\lambda$-modules from the beginning (which also forces the determinant to be trivial and the arithmetic/geometric monodromy groups to coincide, for exponential sums and super-even characters), we will not encounter the difficulty observed in [Kowalski 2008b; Cha et al. 2017] that the normalized characteristic polynomials may be defined over a quadratic extension of the base field, with the possibility of a different Galois group. This was overcome in ibid. by looking at squares of the roots, and showing that their Galois group was still maximal from a study of additive relations, in addition to the multiplicative ones.

**5A.** ***Integral monodromy groups.*** The lisse sheaves $\mathcal{F}_\lambda$ of free modules on a variety $X$ given by Theorems 2.1, 3.2 and 4.2 correspond to continuous representations $\rho_\lambda : \pi_1(X, \bar{\eta}) \to \mathrm{GL}_r(\mathcal{O}_\lambda)$, for $\bar{\eta}$ a geometric generic point, such that, for every $x \in X(\mathbb{F}_q)$, if $\mathrm{Frob}_{x,q} \in \pi_1(X, \bar{\eta})^\sharp$ is the geometric Frobenius conjugacy class at $x$, then $\rho_\lambda(\mathrm{Frob}_{x,q}) \in \mathrm{GL}_r(\mathcal{O}_\lambda)^\sharp$ gives the action of $\mathrm{Frob}_q$ on $(\mathcal{F}_\lambda)_x$.

**Definition 5.4** (monodromy groups). The *geometric and arithmetic monodromy groups* of $\rho_\lambda$ are respectively

$$G_\lambda^{\mathrm{geom}} := \overline{\rho_\lambda(\pi_1^{\mathrm{geom}}(X, \bar{\eta}))}^{\mathrm{Zar}} \leq G_\lambda := \overline{\rho_\lambda(\pi_1(X, \bar{\eta}))}^{\mathrm{Zar}} \leq \mathrm{GL}_r(\overline{E_\lambda}),$$

where $\overline{\cdot}^{\mathrm{Zar}}$ denotes Zariski closure in $\mathrm{GL}_r(\overline{E_\lambda})$. By reducing modulo $\lambda$, we also obtain representations $\tilde{\rho}_\lambda : \pi_1(X, \bar{\eta}) \to \mathrm{GL}_r(\mathbb{F}_\lambda)$, and we define the *geometric and arithmetic integral monodromy groups* of $\rho_\lambda$ as the monodromy groups

$$\widetilde{G}_\lambda^{\mathrm{geom}} := \tilde{\rho}_\lambda(\pi_1^{\mathrm{geom}}(X, \bar{\eta})) \leq \widetilde{G}_\lambda := \tilde{\rho}_\lambda(\pi_1(X, \bar{\eta})) \leq \mathrm{GL}_r(\mathbb{F}_\lambda)$$

of $\tilde{\rho}_\lambda$. If the adjective "projective" is added to those groups, one refers to their image with respect to the projections $\mathrm{GL}_r \to \mathrm{PGL}_r$ (over $\overline{E_\lambda}$ or $\mathbb{F}_\lambda$ respectively).

**5A1.** *From monodromy to integral monodromy.* The determination of integral monodromy groups may be more challenging that their counterparts over $\overline{E_\lambda}$, since they have less structure (under purity assumption, the connected component at the identity of $G_\lambda^{\mathrm{geom}}$ is a semisimple algebraic group).

Fortunately, as explained by [Katz 2012b, Section 7], one may use deep results of [Larsen and Pink 1992; Larsen 1995] to conclude (roughly) that if the monodromy over $\overline{E_\lambda}$ is as large as possible, then the same holds for a density 1 of the integral monodromy groups.

Katz's argument is given for sheaves of $\mathbb{Z}_\ell$-modules, but carries over more generally to sheaves of $\mathcal{O}_\lambda$-modules: we spelled out the details in [Perret-Gentil 2018a, Section 5.2], and the conclusion reads as:

**Theorem 5.5.** *Let $X$ be a smooth affine geometrically connected variety over $\mathbb{F}_p$, let $E \subset \mathbb{C}$ be a Galois number field with ring of integers $\mathcal{O}$, and let $\Lambda$ be a set of valuations on $\mathcal{O}$ of natural density* 1. *Let $(\mathcal{F}_\lambda)_{\lambda\in\Lambda}$ be a compatible system with $\mathcal{F}_\lambda$ a lisse sheaf of free $\mathcal{O}_\lambda$-modules on $X$. We assume that*

> *there exists $G \in \{\mathrm{SL}_n, \mathrm{Sp}_{2n}\}$ such that for every $\lambda \in \Lambda$, the arithmetic monodromy group of $\mathcal{F}_\lambda$ is conjugate to $G(\overline{E_\lambda})$.*

*Then there exists a subset $\Lambda_p \subset \Lambda \cap \mathrm{Spec}_{1,p}(\mathcal{O})$ of natural density* 1, *depending on $p$ and on the family, such that $\mathcal{F}_\lambda$ has geometric and arithmetic integral monodromy groups conjugate to $G(\mathbb{F}_\lambda)$ for all $\lambda \in \Lambda_p$.*

**Remark 5.6** (implied constants). The dependency of the sets of valuations on some of the variables $p, k, m$ in Theorems 5.10, 5.11 and 5.12 will give dependencies on those of the implied constants in the final results.

**Remark 5.7** (strong approximation). Another method to get information on integral monodromy groups from the transcendental ones is through strong approximation results for arithmetic groups, as explained in [Katz 2012b, Section 9] (see also [Jouve et al. 2013, Section 5]); this is for example used in [Cha et al. 2017]. In those cases, [Pink 2000] (a generalization of [Matthews et al. 1984; Weisfeiler 1984]) allows to show that the integral monodromy is large for all but finitely many primes. Moreover, by also using results of [Larsen and Pink 1992], it avoids the classification of finite simple groups, unlike [Matthews et al. 1984; Weisfeiler 1984].

However, this requires that the sheaves $\mathcal{F}_\lambda$ on $X$ may be formed over the analytification $X^{\mathrm{an}}$: a sheaf $\mathcal{F}^{\mathrm{an}}$ of finitely generated $\mathcal{O}$-modules is constructed on $X^{\mathrm{an}}$, whose extension of scalars to $\mathcal{O}_\lambda$ corresponds to the analytification of $\mathcal{F}_\lambda$, and strong approximation can then be applied to the monodromy of $\mathcal{F}^{\mathrm{an}}$ in $G(\mathcal{O})$ to yield the result. This can be done in the case of families of $L$-functions considered in [Katz 2012b; Cha et al. 2017], but a priori not for the sheaves from Theorems 2.1 and 3.2 (one may think about Artin–Schreier sheaves, i.e., Kloosterman sheaves of rank 1, as a first example).

**5A2.** *Kloosterman and Birch sheaves.* Combining Theorem 5.5 with the determination of monodromy groups over $\bar{E}_\lambda$ by Katz, we obtain the following:

**Theorem 5.8** (Kloosterman sheaves). *In the setting of Theorem 2.1, there exists a subset $\Lambda_{r,p}$ of $\mathrm{Spec}_{1,p}(\mathcal{O})$, of natural density* 1, *such that for every $\lambda \in \Lambda_{r,p}$, the arithmetic and geometric integral monodromy groups of $\mathcal{K}l_{r,\lambda}$ are equal and conjugate to $\mathrm{SL}_r(\mathbb{F}_\lambda)$ if $r$ is odd, $\mathrm{Sp}_r(\mathbb{F}_\lambda)$ if $r$ is even.*

*Proof.* This follows from Theorem 5.5 and the determination of monodromy groups over $\bar{E}_\lambda$ contained in [Katz 1988, Chapter 11]. □

**Remark 5.9.** By work of Hall [2008] or J-K. Yu (unpublished) when $r = 2$, and the author [Perret-Gentil 2018b] for any $r \geq 2$, one may actually take

$$\Lambda_{r,p} = \{\lambda \in \mathrm{Spec}_{1,p}(\mathcal{O}) \text{ above } \ell : \ell \gg_r 1\}. \tag{21}$$

In particular, the densities of elements $\Lambda_{r,p}$ with bounded norm are bounded from below independently of $p$.

**Theorem 5.10** (Birch sheaves). *In the setting of Theorem 2.1*(2), *there exists a subset* $\Lambda_p$ *of* $\mathrm{Spec}_{1,p}(\mathcal{O})$, *of natural density* 1, *such that for every* $\lambda \in \Lambda_p$, *the arithmetic and geometric integral monodromy groups of* $\mathcal{B}i_\lambda$ *are equal to* $\mathrm{SL}_2(\mathbb{F}_\lambda)$.

*Proof.* This follows from Theorem 5.5 and the determination of monodromy groups over $E_\lambda$ in [Katz 1990, 7.12]. □

**5A3.** *Primitive super-even characters.*

**Theorem 5.11.** *In the setting of Theorem 3.2* (3), *assuming that* $k \geq 4$, *there exists a subset* $\Lambda_{k,p} \subset \mathrm{Spec}_{1,p}(\mathcal{O})$ *of natural density* 1 *such that for every* $\lambda \in \Lambda_{k,p}$, *the arithmetic and geometric integral monodromy groups of* $\mathcal{G}_{k,\lambda}$ *are equal and conjugate to* $\mathrm{Sp}_{2\kappa-2}(\mathbb{F}_\lambda)$.

*Proof.* This follows from Theorem 5.5 and the determination of monodromy groups over $E_\lambda$ in [Katz 2017, Theorem 2.5] (using results from [Katz 2005, 3.10]). □

**5A4.** *Primitive even characters mod* $T^m$.

**Theorem 5.12.** *In the setting of Theorem 4.2* (3), *assuming that* $m \geq 5$ *is odd*, *there exists a subset* $\Lambda_{m,p} \subset \mathrm{Spec}_{1,p}(\mathcal{O})$ *of natural density* 1 *such that for every* $\lambda \in \Lambda_{m,p}$, *the projective arithmetic and geometric integral monodromy groups of* $\mathcal{G}_{m,\lambda}$ *are conjugate to* $\mathrm{PSL}_{m-2}(\mathbb{F}_\lambda)$.

*Proof.* By, [Katz 2013b, Theorem 5.1],

$$\mathrm{SL}_{m-2}(\mathbb{C}) \leq G_{\mathrm{geom}}(\mathcal{G}_{m,\lambda}) \leq G_{\mathrm{arith}}(\mathcal{G}_{m,\lambda}) \leq \mathrm{GL}_{m-2}(\mathbb{C}),$$

whence $PG_{\mathrm{geom}}(\mathcal{G}_{m,\lambda}) = PG_{\mathrm{arith}}(\mathcal{G}_{m,\lambda}) = \mathrm{PGL}_{m-2}(\mathbb{C})$.

However, projective representations are not directly handled in Theorem 5.5. Instead, we note that if $\lambda \in \mathrm{Spec}_{1,p}(\mathcal{O})$ is above $\ell \nmid m-2$, then $\ell \equiv 1 \pmod{m-2}$ (by the characterization of ideals of degree 1 in cyclotomic extensions), so Hensel's lemma implies that every element of $\mathcal{O}_\lambda$ has an $(m-2)$-th root, whence $\mathrm{PGL}_{m-2}(\mathcal{O}_\lambda) \cong \mathrm{SL}_{m-2}(\mathcal{O}_\lambda)$.

If $\mathcal{G}_{m,\lambda}$ corresponds to a representation $\rho_\lambda : \pi_1(X, \bar{\eta}) \to \mathrm{GL}_{m-2}(\mathcal{O}_\lambda)$ and $\pi : \mathrm{GL}_{m-2} \to \mathrm{PGL}_{m-2}$ is the projection, we get in this case a continuous representation $\pi \circ \rho_\lambda : \pi_1(X, \bar{\eta}) \to \mathrm{SL}_{m-2}(\mathcal{O}_\lambda)$ with transcendental arithmetic and geometric monodromy groups isomorphic to $\mathrm{SL}_{m-2}(\mathbb{C})$. We may then apply Theorem 5.5 to get that the arithmetic and geometric integral monodromy of $\pi \circ \rho_\lambda$ are

$\mathrm{SL}_{m-2}(\mathbb{F}_\lambda) \cong \mathrm{PGL}_{m-2}(\mathbb{F}_\lambda)$ for a subset of density 1 of $\mathrm{Spec}_{1,p}(\mathcal{O})$. Since $\mathrm{im}(\pi \circ \rho_\lambda \pmod *\lambda) = \pi(\mathrm{im}\,\rho_\lambda \pmod *\lambda)$, this proves the assertion on the projective monodromy groups of $(\mathcal{G}_{m,\lambda})_\lambda$. □

**5B.** ***Large sieve for Frobenius, with wild ramification.*** Next, we need a version of the large sieve for Frobenius, originally developed in [Kowalski 2006]; see also [Kowalski 2008a; 2008b].

In these works as well as in [Cha et al. 2017], the sieve applies to sheaves of $\mathbb{F}_\ell$-modules on a variety $X$ over $\mathbb{F}_p$, that either

(1) are compatible systems, with $X$ a curve;

(2) are tamely ramified;

(3) have monodromy group of cardinality prime to $p$, a stronger condition than the previous one.

For Kloosterman and Birch sums, (1) applies. However, for super-even characters, the variety is not a curve, and the sheaves are a priori not tamely ramified, which rules out (2). Concerning (3), note that for $E = \mathbb{Q}(\zeta_{p^N})$ and $\lambda \in \mathrm{Spec}(\mathcal{O})$, the prime $p$ always divides $|\mathrm{SL}_r(\mathbb{F}_\lambda)|$ and $|\mathrm{Sp}_r(\mathbb{F}_\lambda)|$ (if $r$ is even).

**5B1.** *Extension of the large sieve for Frobenius.* Instead, we give an extension of [Kowalski 2006, Theorem 3.1; 2008b, Theorems 4.1, 4.3] that works in this case and answers the question in [Kowalski 2006, Remark 4.8]. To bound the sums of Betti numbers that appear, we give two arguments:

(1) Theorem 5.14(b), involving sums of Betti numbers associated to tensor powers of the sheaves, inspired by [Kowalski 2006, Section 4; Katz and Sarnak 1999, Theorem 9.2.6; Katz 2017, Lemma 5.2] and an effective/modular version of a theorem of Burnside on irreducible representations contained in tensor powers of faithful representations.

(2) Theorem 5.14(c), provided by Will Sawin, reducing to the tame case (where a result of Deligne [Illusie 1981] on the Euler characteristic of tamely ramified sheaves can be applied) by exploiting the presence of a compatible system. This gives a much stronger bound, but with less explicit constants.

**Definition 5.13.** Let $X$ be a smooth affine geometrically connected algebraic variety over $\mathbb{F}_p$, $E$ be a number field with ring of integers $\mathcal{O}$, let $\lambda, \lambda' \in \mathrm{Spec}_{1,p}(\mathcal{O})$, and let $\mathcal{F}$ be a lisse sheaf of $R$-modules on $X$, where $R = \overline{\mathbb{Q}}_\ell$, $\mathcal{O}_\lambda$, $\mathcal{O}_\lambda \otimes \mathcal{O}_{\lambda'}$, $\mathbb{F}_\lambda$, or $\mathbb{F}_\lambda \otimes \mathbb{F}_{\lambda'}$. We define the sum of Betti numbers

$$\sigma_c(X, \mathcal{F}) = \sum_{i=0}^{2\dim X} \mathrm{rank}\, H_c^i(X, \mathcal{F}),$$

where the rank of an $R$-module is defined as its dimension over the total ring of fractions of $R$ (recall that these cohomology groups are finitely generated by [SGA 4½ 1977, Exposé 1, Théorème 4.6.2]).

| $G$ | $\dim G$ | $\mathrm{rank}\, G$ | $E_G$ | Type | Weyl group |
|---|---|---|---|---|---|
| $\mathrm{SL}_r$ | $r^2-1$ | $r-1$ | $\frac{1}{2}(2r^2+r-3)$ | $A_{r-1}$ | $\mathfrak{S}_r$ |
| $\mathrm{Sp}_r$ | $\frac{1}{2}(r(r+1))$ | $r/2$ | $\frac{1}{4}(r(2r+3))$ | $C_{r/2}$ | $W_r \leq \mathfrak{S}_r$ |

**Table 1.** Reminder of certain invariants for the groups considered.

If $X$ is a curve and $R = \mathcal{O}_\lambda$, we moreover define

$$\mathrm{cond}(\mathcal{F}_\lambda) = 1 - \chi_c(X, \mathbb{Q}_\ell) + 2\sum_x \mathrm{Swan}_x(\mathcal{F}_\lambda)$$

to be the quantity in [Kowalski 2006, (4.1)] (see also [Katz 1988, Chapters 1–2]), where the sum is over "points at infinity" of $X$.

**Theorem 5.14.** *Let $X$ be a smooth affine geometrically connected algebraic variety of dimension $d$ over $\mathbb{F}_p$. For $E$ a number field with ring of integers $\mathcal{O}$, let $\Lambda \subset \mathrm{Spec}_{1,p}(\mathcal{O})$ with lower density*

$$\delta_\Lambda := \liminf_{L\to\infty} \frac{|\{\lambda \in \Lambda : N(\lambda) \le L\}|}{L/\log L} > 0.$$

*For every $\lambda \in \Lambda$, let $\mathcal{F}_\lambda$ be a rank $r$ lisse sheaf of $\mathbb{F}_\lambda$-modules on $X$, corresponding to a representation*

$$\rho_\lambda : \pi_1(X, \bar\eta) \to \mathrm{GL}_r(\mathbb{F}_\lambda), \tag{22}$$

*for $\bar\eta$ a geometric generic point. We assume that there exists $G \in \{\mathrm{SL}_r, \mathrm{Sp}_r\}$ such that either*

(i) *the arithmetic and geometric monodromy groups of $\rho_\lambda$ are equal and conjugate to $G(\mathbb{F}_\lambda)$ for all $\lambda \in \Lambda$; or*

(ii) *the **projective** arithmetic and geometric monodromy group of $\rho_\lambda$ are equal and conjugate to $\mathrm{PGL}_r(\mathbb{F}_\lambda)$ for all $\lambda \in \Lambda$, and $\zeta_r \in E$, so that $\mathrm{PGL}_r(\mathbb{F}_\lambda) = \mathrm{SL}_r(\mathbb{F}_\lambda) = G(\mathbb{F}_\lambda)$.*[4]

*Let $t \ge 1$ be an integer. For every $\lambda \in \Lambda$, let $\mathbf{\Omega}_\lambda \subset G(\mathbb{F}_\lambda)^t$ be a conjugacy-invariant subset, such that*

$$\delta_{\mathbf{\Omega}} := \sup_{\lambda\in\Lambda} \frac{|\mathbf{\Omega}_\lambda|}{|G(\mathbb{F}_\lambda)|^t} < 1.$$

*Then, for any field $\mathbb{F}_q$ of characteristic $p$ and any $L \ge 1$,*

$$\begin{aligned} P\big(q, (\mathcal{F}_\lambda, \mathbf{\Omega}_\lambda)_{\lambda\in\Lambda}\big) &:= \frac{|\{\boldsymbol{x} \in X(\mathbb{F}_q)^t : (\rho_\lambda(\mathrm{Frob}_{x_i,q}))_i \in \mathbf{\Omega}_\lambda \text{ for all } \lambda \in \Lambda\}|}{|X(\mathbb{F}_q)|^t} \\ &\ll \frac{1}{(1-\delta_{\mathbf{\Omega}})\delta_\Lambda} \frac{\log L}{L}\left(1 + \frac{tC(L, (\mathcal{F}_\lambda)_{\lambda\in\Lambda})^t}{q^{1/2}}\right), \end{aligned}$$

*where*

(a) *if $d = 1$, $C(L, (\mathcal{F}_\lambda)_{\lambda\in\Lambda}) \ll r^{\delta_{G=\mathrm{SL}_r}} L^{\dim G + ((\mathrm{rank}\, G)/2)} \max_{N(\lambda)\le L} \mathrm{cond}(\mathcal{F}_\lambda)$;*

(b) *if $d \ge 1$, $C(L, (\mathcal{F}_\lambda)_{\lambda\in\Lambda}) \ll d L^{\dim G} \max_{N(\lambda)\le L} \max_{M \le N(\lambda)^{M_G}} \sigma_c(X, \mathcal{F}_\lambda^{\otimes M})^2$, with $M_G = \mathrm{rank}(G)(\mathrm{rank}(G)+1)/2$;*

(c) *if the representations (22) arise from a compatible system $\rho : \pi_1(X, \bar\eta) \to \mathrm{GL}_r\big(\prod_{\lambda\in\Lambda} \mathcal{O}_\lambda\big)$, and $X$ has a compactification where it is the complement of a divisor with normal crossing, then*

$$C(L, (\mathcal{F}_\lambda)_{\lambda\in\Lambda}) \ll L^{\dim G + 1} r^{\delta_{G=\mathrm{SL}_r}} (r + C(X, \rho_{\lambda_0})),$$

*where $C(X, \rho_{\lambda_0})$ only depends on $X$ and $\rho_{\lambda_0}$ for an arbitrary fixed $\lambda_0 \in \Lambda$.*

[4] See the proof of Theorem 5.12, recalling that $\lambda$ has degree 1.

**Remarks 5.15.** (1) In the case of curves ($d = 1$) with $E = \mathbb{Q}$ and (i), this is [Kowalski 2006, Theorem 3.1, Proposition 3.3]; see also [Kowalski 2008b, Section 5, Remark 5.4].

(2) We handle the weaker (ii) on projective monodromy groups to treat $L$-function attached to even Dirichlet characters over function fields (Section 4).

(3) The constant $C(L, (\mathcal{F}_\lambda)_{\lambda\in\Lambda})$ may depend on the characteristic $p$, but crucially not on the index $[\mathbb{F}_q : \mathbb{F}_p]$.

(4) The last part of [Kowalski 2006, Remark 5.2] does not seem quite correct: one crucially has to control the dependency of $C$ with respect to $L$ (that is, the Betti numbers) if one wants to take $L \to \infty$.

In practice, we will use the following consequence of Theorem 5.14:

**Corollary 5.16.** *In the setting of Theorem 5.14*:

(a) *If $X$ is a curve, then*

$$P\big(q, (\mathcal{F}_\lambda, \boldsymbol{\Omega}_\lambda)_{\lambda\in\Lambda}\big) \ll \frac{t \sup_{\lambda\in\Lambda} \operatorname{cond}(\mathcal{F}_\lambda)^t}{(1-\delta_{\boldsymbol{\Omega}})\delta_\Lambda} \frac{\log q}{q^{1/(tE_G)}},$$

*where the implied constant is absolute and $E_G = \dim G + (\operatorname{rank} G)/2$.*

(b) *If there are constants $B_1 > 0$ and $B_2 > 1$ such that*

$$\sup_{\lambda\in\Lambda} \sigma_c(X, \mathcal{F}_\lambda^{\otimes N}) \le B_1 B_2^N \qquad \text{for all } N \ge 1,$$

*then*

$$P\big(q, (\mathcal{F}_\lambda, \boldsymbol{\Omega}_\lambda)_{\lambda\in\Lambda}\big) \ll \frac{t^2 (B_1^2 d r^{\delta_{G=\mathrm{SL}_r}})^t (\log(B_2) M_G + \dim G)}{(1-\delta_{\boldsymbol{\Omega}})\delta_\Lambda} \frac{\log\log q}{\log q},$$

*with an absolute implied constant.*

(c) *If hypothesis (c) of Theorem 5.14 holds, then*

$$P\big(q, (\mathcal{F}_\lambda, \boldsymbol{\Omega}_\lambda)_{\lambda\in\Lambda}\big) \ll \frac{t (r^{\delta_{G=\mathrm{SL}_r}+1} C(X, \rho_{\lambda_0}))^t}{(1-\delta_{\boldsymbol{\Omega}})\delta_\Lambda} \frac{\log q}{q^{1/(2t(\dim G+1))}}.$$

We prove the theorem and its corollary in the next sections.

### 5C. *Preliminaries to the proof of Theorem 5.14(b).*

**5C1.** *Irreducibles in tensor powers of faithful representations.* A classical theorem of Burnside asserts that

> if $G$ is a finite group with a faithful (complex) representation $\rho$, then any irreducible representation of $G$ appears as a direct summand of $\rho^{\otimes M}$ for some integer $M \ge 1$

(see, e.g., [Steinberg 1962; Brauer 1964; Bryant and Kovács 1972]). The same result holds for compact groups, and is the key to get bounds on Betti numbers in [Katz 2017]. For classical groups, this can actually directly be seen from Weyl's constructions of the irreducible modules.

A key input to the proof of Theorem 5.14(b) is the following modular version of Burnside's result, for classical finite groups in defining characteristic.

**Proposition 5.17.** *Let $k$ be a field of characteristic $\ell$ and $G = \mathrm{SL}_n(\mathbb{F}_\ell)$ or $\mathrm{Sp}_n(\mathbb{F}_\ell)$ with its standard $k$-representation* $\mathrm{Std} : G \to \mathrm{GL}_n(k)$. *Any irreducible $k$-representation of $G$ appears as a composition factor*[5] *of* $\mathrm{Std}^{\otimes M}$ *for some $M \leq \ell M_G$, $M_G = \mathrm{rank}(G)(\mathrm{rank}(G)+1)/2$. Therefore, for any $k$-representation $\pi$ of $G$, the semisimplification $\pi^{\mathrm{ss}}$ appears as a direct summand of* $(\dim \pi)(\mathrm{Std}^{\otimes M})^{\mathrm{ss}}$.

*Proof.* Since $G$ is defined over $\mathbb{F}_\ell$, any irreducible $k$-representation of $G$ is absolutely irreducible, because $\mathbb{F}_\ell$ is the splitting field of $G$ by a 1968 result of Steinberg [Humphreys 2006, Section 5.2].

By a 1963 lifting theorem of Steinberg (see [Humphreys 2006, Section 2.11]), the absolutely irreducible representations of $G$ in characteristic $\ell$ are given by the modules $L(\lambda)$ with $\lambda$ an $\ell$-restricted highest weight, i.e., $0 \leq \langle \lambda, \alpha^\vee \rangle < q$ for all $\alpha \in \Delta$. For $\omega_i$ $(1 \leq i \leq \mathrm{rank}(G))$ the fundamental dominant weights, that means that $\lambda = \sum_{i=1}^{\mathrm{rank}(G)} a_i \omega_i$ with $0 \leq a_i < \ell$.

In Bourbaki numbering [2005, Tables], $\omega_i$ is $\Lambda^i(\mathrm{Std})$ (see [ibid., VIII.13.1.IV]) (resp. $\ker(\Lambda^i(\mathrm{Std}) \to \Lambda^{i-2}(\mathrm{Std}))$; see [ibid, VIII.13.3.IV]) for $\mathrm{SL}_n$ (resp. $\mathrm{Sp}_n$). These are simple quotients or subrepresentations of $\mathrm{Std}^{\otimes i}$, so they appear in the composition series. □

**Remark 5.18.** For complex representations, combining David Speyer's proof of Burnside's theorem [Speyer 2011] with character bounds [Gluck 1993] shows that $M \ll \dim G$ is enough, as $\ell \to \infty$. Such an improvement (or even $M \ll \log |\mathbb{F}_\ell|$) to Proposition 5.17 would lead to bounds of the quality of Corollary 5.16(c) in Corollary 5.16(b). However, while Brauer characters control composition factors, they do not satisfy (in defining characteristic) good bounds, to extend this characteristic 0 idea.

**5C2.** *Betti numbers of reductions modulo $\lambda$ and semisimplifications.*

**Lemma 5.19.** *In the setting of Theorem 5.14, if $\mathcal{F}_\lambda$ is the sheaf of $\mathbb{F}_\lambda$-modules on $X$ obtained by reduction of a lisse sheaf of $\mathcal{O}_\lambda$-modules $\widehat{\mathcal{F}}_\lambda$ on $X$, then*

$$\sigma_c(X, \widehat{\mathcal{F}}_\lambda^{\otimes M}) \leq \sigma_c(X, \mathcal{F}_\lambda^{\otimes M}) \leq 2\sigma_c(X, \widehat{\mathcal{F}}_\lambda^{\otimes M})$$

*for any $M \geq 1$.*

*Proof.* Let $\mathcal{G} = \mathcal{F}_\lambda^{\otimes M}$ and $\widehat{\mathcal{G}}_\lambda = \widehat{\mathcal{F}}_\lambda^{\otimes M}$. The lower bound appears in [Katz and Sarnak 1999, p. 279], and the same argument yields the upper bound: we have the universal coefficients short exact sequence

$$0 \to H_c^i(X, \widehat{\mathcal{G}}_\lambda) \otimes_{\mathcal{O}_\lambda} \mathbb{F}_\lambda \to H_c^i(X, \mathcal{G}_\lambda) \otimes_{\mathcal{O}_\lambda} \mathbb{F}_\lambda \to H_c^{i+1}(X, \widehat{\mathcal{G}}_\lambda)[\lambda] \to 0,$$

obtained after truncating the long exact sequence in cohomology [SGA 4½ 1977, 1.6.5] associated to the short exact sequence $0 \to \widehat{\mathcal{F}}_\lambda \xrightarrow{\cdot\lambda} \widehat{\mathcal{F}}_\lambda \to \mathcal{F}_\lambda \to 0$ . Taking dimensions, this implies that

$$\sigma_c(X, \widehat{\mathcal{G}}_\lambda) \leq \sigma_c(X, \mathcal{G}_\lambda) \leq \sum_{i \geq 0} \big(\dim H_c^i(X, \hat{\mathcal{G}}_\lambda) + \dim H_c^{i+1}(X, \widehat{\mathcal{G}}_\lambda)\big). \qquad \square$$

[5] We need to look at composition factors instead of summands, since we consider modular representations, which are not completely reducible.

**Remark 5.20.** If the sheaves $\mathcal{F}_\lambda$ in Theorem 5.14 are obtained by reduction of sheaves of $\mathcal{O}_\lambda$-modules $\widehat{\mathcal{F}}_\lambda$, Lemma 5.19 shows that it suffices to check hypothesis in (b) of Corollary 5.16 for $\mathcal{F}_\lambda$, up to replacing $B_1$ by $2B_1$.

To deal with noncompletely reducible representations, we observe the following:

**Lemma 5.21.** *Let $\mathcal{F}$ be a sheaf of $\mathbb{F}_\ell$-modules on $X$ with composition series*

$$0 = \mathcal{F}_0 \subset \cdots \subset \mathcal{F}_n = \mathcal{F}, \quad \mathcal{G}_i := \mathcal{F}_{i+1}/\mathcal{F}_i \text{ simple} \quad (0 \le i \le n-1).$$

*Then $\sigma_c(X, \mathcal{F}^{\mathrm{ss}}) = \sum_{i=0}^{n-1} \sigma_c(X, \mathcal{G}_i) = \sigma_c(X, \mathcal{F})$.*

*Proof.* For all $0 \le i \le n-1$, we have a short exact sequence $0 \to \mathcal{F}_i \to \mathcal{F}_{i+1} \to \mathcal{G}_i \to 0$, which gives for all $a \ge 0$ a long exact sequence in cohomology

$$\cdots \to H_c^a(X, \mathcal{F}_i) \to H_c^a(X, \mathcal{F}_{i+1}) \to H_c^a(X, \mathcal{G}_i) \to H_c^{a+1}(X, \mathcal{F}_i) \to \cdots$$

that yields $\sigma_c(X, \mathcal{F}_{i+1}) = \sigma_c(X, \mathcal{G}_i) + \sigma_c(X, \mathcal{F}_i)$, whence $\sigma_c(X, \mathcal{F}) = \sigma_c(X, \mathcal{F}_n) = \sum_{i=0}^{n-1} \sigma_c(X, \mathcal{G}_i) = \sigma_c(X, \mathcal{F}^{\mathrm{ss}})$. □

**5D.** ***Proof of Theorem 5.14.*** We first give the proof under (i), before indicating the changes required in the projective case (ii).

For $\lambda, \lambda' \in \Lambda$, we will denote by $\ell, \ell'$ the primes above which they respectively lie. Since $\Lambda \subset \mathrm{Spec}_{1,p}(\mathcal{O})$, note that $\mathbb{F}_\lambda = \mathbb{F}_\ell$, $\mathbb{F}_{\lambda'} = \mathbb{F}_{\ell'}$. We also let $\widehat{G(\mathbb{F}_\ell)}$ be the set of irreducible (complex) representations of $G(\mathbb{F}_\ell)$.

For every $\lambda \in \Lambda$, we consider the lisse sheaf $\mathcal{G}_\lambda = \mathcal{F}_\lambda^{\boxtimes t}$ on $X^t$. By [Kowalski 2008b, Lemma 5.1], the natural map $\pi_1(X^t, (\bar\eta, \dots, \bar\eta)) \to \pi_1(X, \bar\eta)^t$ is surjective, so that the arithmetic and geometric monodromy groups of $\mathcal{G}_\lambda$ are equal and conjugate to $G(\mathbb{F}_\lambda)^t$.

Exactly as in [Kowalski 2006, Theorem 3.1, Proposition 3.3, Section 5], we get that

$$P\big(q, (\mathcal{F}_\lambda, \boldsymbol{\Omega}_\lambda)_{\lambda\in\Lambda}\big) \ll \Delta\Bigg[\sum_{\substack{\lambda\in\Lambda\\ N(\lambda)\le L}} \left(1 - \frac{|\boldsymbol{\Omega}_\lambda|}{|G(\mathbb{F}_\lambda)|}\right)\Bigg]^{-1} \ll \frac{\Delta \log L}{\delta_\Lambda(1-\delta_{\boldsymbol{\Omega}})L},$$

where $\Delta \ll 1 + q^{-1/2}\tilde{C}(L, (\mathcal{F}_\lambda)_{\lambda\in\Lambda})$, and $\tilde{C}(L, (\mathcal{F}_\lambda)_{\lambda\in\Lambda})$ is defined by

$$\max_{\substack{\lambda\in\Lambda\\ N(\lambda)\le L}} \max_{\substack{\boldsymbol{\pi}\in\widehat{G(\mathbb{F}_\lambda)^t}\\ \boldsymbol{\pi}\ne 1}} \Bigg[\sum_{\substack{\boldsymbol{\pi}'\in\widehat{G(\mathbb{F}_\lambda)^t}\\ \boldsymbol{\pi}'\ne 1}} \sigma_c(X^t, \mathcal{F}_{\boldsymbol{\pi},\boldsymbol{\pi}'}) + \sum_{\substack{\lambda'\in\Lambda\\ N(\lambda')\le L\\ \ell'\ne\ell}} \sum_{\substack{\boldsymbol{\pi}'\in\widehat{G(\mathbb{F}_{\lambda'})^t}\\ \boldsymbol{\pi}'\ne 1}} \sigma_c(X^t, \mathcal{F}_{\boldsymbol{\pi},\boldsymbol{\pi}'})\Bigg],$$

with (see [Kowalski 2006, Proof of Proposition 5.1])

$$\mathcal{F}_{\boldsymbol{\pi},\boldsymbol{\pi}'} = \tau_{\boldsymbol{\pi},\boldsymbol{\pi}'} \circ \begin{cases} \rho_\lambda^{\boxtimes t} & \text{for } \ell = \ell', \\ (\rho_\lambda^{\boxtimes t}, \rho_{\lambda'}^{\boxtimes t}) & \text{for } \ell \ne \ell', \end{cases} \qquad \tau_{\boldsymbol{\pi},\boldsymbol{\pi}'} = \begin{cases} \boldsymbol{\pi} \otimes D(\boldsymbol{\pi}') & \text{for } \ell = \ell', \\ \boldsymbol{\pi} \boxtimes D(\boldsymbol{\pi}') & \text{for } \ell \ne \ell', \end{cases}$$

identifying lisse sheaves of $\overline{\mathbb{Q}}_\ell$-modules on $X^t$ and continuous representations $\pi_1(X^t, \bar\eta) \to \mathrm{GL}_m(\overline{\mathbb{Q}}_\ell)$. Note that $\rho_\lambda$ and $(\rho_\lambda, \rho_{\lambda'})$ respectively correspond to sheaves of $\mathbb{F}_\ell$- and $\mathbb{Z}/\ell\ell'$-modules (if $\ell \ne \ell'$).

Hence, we need to show that

$$\widetilde{C}(L, (\mathcal{F}_\lambda)_{\lambda\in\Lambda}) \ll tC(L, (\mathcal{F}_\lambda)_{\lambda\in\Lambda})^t,$$

with $C$ defined in the statement of the theorem. Künneth's formula [SGA 4½ 1977, Exposé 6, 2.4] reduces this to the case $t = 1$.

**5D1.** *Case (a): curves.* The first bound on $C(L, (\mathcal{F}_\lambda)_{\lambda\in\Lambda})$ in Theorem 5.14, when $d = 1$, is contained in [Kowalski 2006] (with a power of $L$ smaller by one here, because we assume that the arithmetic and geometric monodromy groups coincide).

**5D2.** *Case (c): compatible systems on varieties by reduction to the tame case.* Let $\lambda_0 \in \Lambda$ be fixed and let $\varphi : Y \to X$ be the étale covering corresponding to $f \pmod{\lambda_0}$. As in [Kowalski 2006, Proposition 4.7], by the Hochschild–Serre sequence,

$$\sigma_c(X, \mathcal{F}_{\pi,\pi'}) \le \sigma_c(Y, \varphi^*\mathcal{F}_{\pi,\pi'}).$$

It then suffices to show that the compatible system $\rho$ is tame when restricted to $Y$. Indeed, a result of Deligne [Illusie 1981, Corollaire 2.8] shows that the Euler characteristic of a lisse tame sheaf is equal to its rank times the Euler characteristic of the variety, so by [Katz 2001, $\sigma - \chi$ inequality, p. 40], we have in this case

$$\begin{aligned}\sigma_c(Y, \varphi^*\mathcal{F}_{\pi,\pi'}) &\ll r + |\chi_c(Y, \varphi^*\mathcal{F}_{\pi,\pi'})| + \sum_{j=1}^{\dim X} |\chi_c(\text{codim } j \text{ in } Y, \varphi^*\mathcal{F}_{\pi,\pi'})| \\ &\le r + \dim(\pi)\dim(\pi')C(X, \rho_{\lambda_0}),\end{aligned}$$

where $C(X, \rho_{\lambda_0})$ is a constant depending only on the Euler characteristics $\chi_c$ of $Y$ and its subvarieties, hence only on $X$ and $\mathcal{F}_{\lambda_0}$. Therefore,

$$\begin{aligned}\widetilde{C}(L, (\mathcal{F}_\lambda)_{\lambda\in\Lambda}) &\ll C(X, \rho_{\lambda_0}) \cdot r \max_{\substack{\lambda\in\Lambda\\ N(\lambda)\le L}} \max_{\substack{\pi\in\widehat{G(\mathbb{F}_\lambda)}\\ \pi\ne 1}} d_\pi \Bigg[ \sum_{\substack{\pi'\in\widehat{G(\mathbb{F}_\lambda)}\\ \pi'\ne 1}} d_{\pi'} + \sum_{\substack{\lambda'\in\Lambda\\ N(\lambda')\le L\\ \ell'\ne\ell}} \sum_{\substack{\pi'\in\widehat{G(\mathbb{F}_{\lambda'})}\\ \pi'\ne 1}} d_{\pi'} \Bigg] \\ &\ll C(X, \rho_{\lambda_0}) \cdot r^{\delta_{G=\mathrm{SL}_r}+1} L^{\dim G+1},\end{aligned}$$

where $d_\pi := \dim \pi$. Indeed, the number (complex) of irreducible representations of $G(\mathbb{F}_\ell)$ is given by $|G(\mathbb{F}_\ell)^\sharp| \ll |Z(G(\mathbb{F}_\ell))|\ell^{\operatorname{rank} G} \le r^{\delta_{G=\mathrm{SL}_r}}\ell^{\operatorname{rank} G}$ (see [Malle and Testerman 2011, Corollary 26.10]), and the maximal dimension of such a representation is $\ll \ell^{\frac{1}{2}(\dim G - \operatorname{rank} G)}$; see [Kowalski 2008a, Proposition 5.4].

To show the tameness of the compatible system restricted to $Y$, first note that it is tame at $\lambda_0$, since it factors by construction through the pro-$\ell_0$-group $\{g \in \mathrm{GL}_n(\mathcal{O}_{\lambda_0}) : g \equiv 1 \pmod{\lambda_0}\}$, where $\ell_0$ is the prime above which $\lambda_0$ lies. By purity, it suffices to look at restriction to curves; see [Illusie 1981, Section 2.6; Kerz and Schmidt 2010]. In this case, [Katz 2002, 7.5.1] shows, from a compatibility result of Deligne, that tameness at one prime implies tameness of the whole system.

**5D3.** *Case (b): varieties through modular representations.* Given $\pi \in \widehat{G(\mathbb{F}_\lambda)}$, $\pi' \in \widehat{G(\mathbb{F}_{\lambda'})}$, we need to bound the sums of Betti numbers $\sigma_c(X, \mathcal{F}_{\pi,\pi'})$. By [Curtis and Reiner 1962, Corollary 75.4], $\pi$ (resp. $\pi'$) is defined over the ring of integers of a finite extension $F_\lambda/E_\lambda$ (resp. $F_{\lambda'}/E_{\lambda'}$), say

$$\pi : G(\mathbb{F}_\lambda) \to \mathrm{GL}_m(\mathcal{O}_{F_\lambda}), \qquad \pi' : G(\mathbb{F}_{\lambda'}) \to \mathrm{GL}_{m'}(\mathcal{O}_{F_{\lambda'}}).$$

By reduction, we obtain

$$\widetilde{\pi} : G(\mathbb{F}_\lambda) \to \mathrm{GL}_m(k), \qquad \widetilde{\pi}' : G(\mathbb{F}_{\lambda'}) \to \mathrm{GL}_{m'}(k'),$$

for the residue field $k/\mathbb{F}_\lambda$, (resp. $k'/\mathbb{F}_{\lambda'}$). Let $\mathrm{Std}_\lambda : G(\mathbb{F}_\lambda) \to \mathrm{GL}_r(\mathbb{F}_\lambda)$ be the standard representation by inclusion.

We start with the case $\ell = \ell'$, which is easier. We may then assume that $\mathbb{F}_\lambda = \mathbb{F}_{\lambda'}$. By Lemmas 5.19 and 5.21, along with the fact that $\rho_\lambda : \pi_1(X, \bar{\eta}) \to G(\mathbb{F}_\lambda)$ is surjective,

$$\sigma_c(X, \mathcal{F}_{\pi,\pi'}) \le \sigma_c(X, \mathcal{F}_{\widetilde{\pi},\widetilde{\pi}'}) \le \sigma_c(X, \tau^{\mathrm{ss}}_{\widetilde{\pi},\widetilde{\pi}'} \circ \rho_\lambda).$$

By Proposition 5.17, every simple summand of $\tau^{\mathrm{ss}}_{\widetilde{\pi},\widetilde{\pi}'}$ appears as a composition factor of $(\mathrm{Std}_\lambda \otimes \overline{\mathbb{F}}_\ell)^{\otimes M}$ for some $M \le \ell M_G$. It follows that

$$\sigma_c(X, \mathcal{F}_{\pi,\pi'}) \le (\dim \pi)(\dim \pi') \max_{M \le \ell M_G} \sigma_c(X, \mathcal{F}_\lambda^{\otimes M}). \tag{23}$$

Let us now assume that $\ell \ne \ell'$, and note that $(\rho_\lambda, \rho_{\lambda'})$ corresponds to the sheaf of $\mathbb{Z}/\ell\ell'$-modules on $X$ given by $\Delta^*(\mathcal{F}_\lambda \boxtimes \mathcal{F}_{\lambda'})$, for $\Delta : X \to X \times X$ the diagonal immersion. We may view $\mathcal{F}_{\pi,\pi'}$ as sheaf of $(\mathcal{O}_{F_\lambda} \otimes \mathcal{O}_{F_{\lambda'}})$-modules, and $\sigma_c(X, \mathcal{F}_{\pi,\pi'})$ is equal to the sum of the ranks (under Definition 5.13) of the corresponding étale cohomology groups with compact support. Then $\mathcal{F}_{\widetilde{\pi},\widetilde{\pi}'}$ is a sheaf of $(k \otimes k')$-modules, and by Lemma 5.21 and the same argument as in Lemma 5.19,

$$\sigma_c(X, \mathcal{F}_{\pi,\pi'}) \le \sigma_c(X, \mathcal{F}_{\widetilde{\pi},\widetilde{\pi}'}) = \sigma_c\big(X, \tau^{\mathrm{ss}}_{\widetilde{\pi},\widetilde{\pi}'} \circ \Delta^*(\mathcal{F}_\lambda \boxtimes \mathcal{F}_{\lambda'})\big).$$

As above, we get that every simple summand in $\tau^{\mathrm{ss}}_{\widetilde{\pi},\widetilde{\pi}'}$ appears as a composition factor of the $(k \otimes k')$-module $\mathrm{Std}^{\otimes M} \boxtimes \mathrm{Std}^{\otimes M'}$ for some $M \le \ell M_G$ and $M' \le \ell' M_G$. This implies that

$$\sigma_c(X, \mathcal{F}_{\widetilde{\pi},\widetilde{\pi}'}) \le (\dim \pi + \dim \pi') \max_{M \le \ell M_G} \max_{M' \le \ell' M_G} \sigma_c(X, \Delta^* \mathcal{G}_{M,M'}),$$

where $\mathcal{G}_{M,M'} = (\mathcal{F}_\lambda \otimes k)^{\otimes M} \boxtimes (\mathcal{F}_{\lambda'} \otimes k')^{\otimes M'}$.

By purity [Fu 2011, Corollary 8.5.6] and the localization sequence [Fu 2011, Proposition 5.6.11], this implies that $\sigma_c(X, \Delta^* \mathcal{G}_{M,M'}) \le \sigma_c(X \times X, \mathcal{G}_{M,M'})$. By Künneth's formula [SGA 4½ 1977, Exposé 6, 2.4],

$$\operatorname{rank} H^i_c(X \times X, \mathcal{G}_{M,M'}) = \sum_{a+b=i} \operatorname{rank} H^a_c(X, \mathcal{F}_\lambda^{\otimes M}) \operatorname{rank} H^b_c(X, \mathcal{F}_{\lambda'}^{\otimes M'}) \le \sigma_c(X, \mathcal{F}_\lambda^{\otimes M}) \sigma_c(X, \mathcal{F}_{\lambda'}^{\otimes M'}),$$

hence

$$\sigma_c(X, \mathcal{F}_{\pi,\pi'}) \ll d(\dim \pi + \dim \pi') S(\lambda) S(\lambda') \tag{24}$$

where $S(\lambda) := \max_{M \le N(\lambda) M_G} \sigma_c(X, \mathcal{F}_\lambda^{\otimes M})$.

Thus, (23) and (24) yield that, as in Section 5D2,

$$\widetilde{C}(L,(\mathcal{F}_\lambda)_{\lambda\in\Lambda}) \ll \max_{\substack{\lambda\in\Lambda\\ N(\lambda)\le L}} S(\lambda) \max_{\substack{\pi\in\widehat{G(\mathbb{F}_\ell)}\\ \pi\neq 1}} \Bigg[ d_\pi \sum_{\substack{\pi'\in\widehat{G(\mathbb{F}_\ell)}\\ \pi'\neq 1}} d_{\pi'} + d \sum_{\substack{\lambda'\in\Lambda\\ N(\lambda')\le L\\ \ell'\neq\ell}} \sum_{\substack{\pi'\in\widehat{G(\mathbb{F}_{\ell'})}\\ \pi'\neq 1}} (d_\pi + d_{\pi'}) S(\lambda') \Bigg]$$

$$\ll d r^{\delta_{G=\mathrm{SL}_r}} L^{\dim G} \max_{N(\lambda)\le L} S(\lambda)^2.$$

**5D4.** *Projective monodromy groups.* Let us now suppose that only (ii) holds. For $\eta : G \to PG$ the projection, we have

$$P\big(q,(\mathcal{F}_\lambda,\boldsymbol{\Omega}_\lambda)_{\lambda\in\Lambda}\big) \le \frac{|\{\boldsymbol{x}\in X(\mathbb{F}_q)^t : (\eta\rho_\lambda(\mathrm{Frob}_{x_i,q}))_i \in \eta(\boldsymbol{\Omega}_\lambda) \text{ for all } \lambda\in\Lambda\}|}{|X(\mathbb{F}_q)|^t},$$

and for any $\Omega \subset G(\mathbb{F}_\lambda)$,

$$\frac{|\eta(\Omega)|}{|PG(\mathbb{F}_\lambda)|} = \frac{|\eta(\Omega)||Z(G(\mathbb{F}_\lambda))|}{|G(\mathbb{F}_\lambda)|} \le \frac{|\Omega|}{|G(\mathbb{F}_\lambda)|} = \delta_\Omega.$$

Thus, it is enough to repeat the arguments above with $G$ replaced by $PG$. Indeed, since $(r, |\mathbb{F}_\lambda|-1) = r$ for all $\lambda\in\Lambda$, we have $\mathrm{SL}_r(\mathbb{F}_\lambda) \cong \mathrm{PGL}_r(\mathbb{F}_\lambda)$, so this can be done mutatis mutandis (in particular, the "standard representation" $PG(\mathbb{F}_\lambda) \to \mathrm{GL}_r(\mathbb{F}_\lambda)$ on page 1321 is well-defined). □

## 6. Generic maximality of splitting fields and linear independence

This section mostly recalls some results from [Kowalski 2008b] and gives their analogues for SL when necessary.

### 6A. *Generic maximality of splitting fields.*

**Definition 6.1.** For $R$ a ring and $r \ge 2$ an integer, we let

$$\mathcal{P}_{\mathrm{SL}_r}(R) := \{P\in R[T] \text{ monic} : \deg(P) = r,\ P(0) = 1\} \qquad (r\ge 2),$$

$$\mathcal{P}_{\mathrm{Sp}_r}(R) := \{P\in \mathcal{P}_{\mathrm{SL}_r}(R) : P(T) = T^r P(1/T)\} \qquad (r\ge 2 \text{ even}).$$

Note that for $G\in\{\mathrm{SL}_r, \mathrm{Sp}_r\}$, the set of (reversed) characteristic polynomials of elements of $G(R)$ is included in $\mathcal{P}_G(R)$, with equality at least when $R$ is a finite field; see the reference to Chavdarov's proof in [Kowalski 2008a, Lemma B.5(2)].

Let $E$ be a Galois number field with ring of integers $\mathcal{O}$. Note that the Galois group of a polynomial $P\in\mathcal{P}_G(\bar{E})$ of degree $n$ is contained in

- $\mathfrak{S}_r$ if $G = \mathrm{SL}_r$.
- $W_r \le \mathfrak{S}_r$ (the Coxeter group $B_{r/2}$) if $G = \mathrm{Sp}_r$ ($r$ even).

We will say that the Galois group is *nonmaximal* if this inclusion is strict.

**6A1.** *Detecting nonmaximal Galois groups.*

**Proposition 6.2.** *Let $G = \mathrm{SL}_r$ ($r \geq 2$) or $G = \mathrm{Sp}_r$ ($r \geq 2$ even). For every $t \geq 1$ and $\lambda \in \mathrm{Spec}_1(\mathcal{O})$, there exist conjugacy-invariant sets $\boldsymbol{\Omega}_{\boldsymbol{i},\lambda,G^t} \subset G(\mathbb{F}_\lambda)^t$ ($\boldsymbol{i} \in \boldsymbol{I}$, with $\boldsymbol{I}$ an index set of size $4t$) such that*:

- *$\boldsymbol{\Omega}_{i,\lambda,G^t}$ has density $\leq \delta_{r,t} := ((1-(1/r!))(1+(r/\ell)))^t(1-(1/2r))$.*
- *If $\boldsymbol{g} = (g_1, \dots, g_t) \in \mathcal{P}_G(\mathcal{O}_\lambda)^t$ is such that $\prod_{i=1}^t \det(1 - Tg_i) \in \mathcal{P}_G(\mathcal{O}_\lambda) \subset \mathcal{O}_\lambda[T]$ has nonmaximal Galois group, that is, strictly contained in $\mathfrak{S}_r^t$ (resp. $W_r^t$) if $G = \mathrm{SL}_r$ (resp. $\mathrm{Sp}_r$), then there exists $\boldsymbol{i} \in \boldsymbol{I}$ such that $\boldsymbol{g} \pmod{*}\lambda \in \boldsymbol{\Omega}_{\boldsymbol{i},\lambda,G^t}$.*

*Proof.* The case $G = \mathrm{Sp}_r$ is contained in [Kowalski 2008b, Proof of Theorem 4.3] (see also [Kowalski 2008a, Proof of Theorem 8.13]), using [Kowalski 2008a, Lemma B.5] to switch between densities of matrices and characteristic polynomials, and up to replacing $\mathbb{Z}$ by $\mathcal{O}_\lambda$.

The case $G = \mathrm{SL}_r$ is simpler, and we also apply the lemma of Bauer quoted by Gallagher [1973, p. 98]: if $H \leq \mathfrak{S}_r$ is transitive, contains a transposition and a $m$-cycle with $m > r/2$ prime, then $H = \mathfrak{S}_r$. We define

$$\begin{aligned}
\widetilde{\Omega}_{0,\lambda} &= \{P \in \mathcal{P}_{\mathrm{SL}_r}(\mathbb{F}_\lambda) : \text{product of linear factors}\}^c,\\
\widetilde{\Omega}_{1,\lambda} &= \{P \in \mathcal{P}_{\mathrm{SL}_r}(\mathbb{F}_\lambda) : P \text{ reducible}\},\\
\widetilde{\Omega}_{2,\lambda} &= \{P \in \mathcal{P}_{\mathrm{SL}_r}(\mathbb{F}_\lambda) : P = QQ_1\cdots Q_s,\ Q, Q_i \text{ irreducible},\ \deg(Q) = 2,\ \deg(Q_i) \text{ odd}\}\\
\widetilde{\Omega}_{3,\lambda} &= \{P \in \mathcal{P}_{\mathrm{SL}_r}(\mathbb{F}_\lambda) : P \text{ has an irreducible factor of prime degree} > r/2\}^c,\\
\Omega_{j,\lambda} &= \{g \in \mathrm{SL}_r(\mathbb{F}_\lambda) : \det(1-Tg) \in \widetilde{\Omega}_{j,\lambda}\} \qquad (0 \leq j \leq 3),\\
\boldsymbol{\Omega}_{\boldsymbol{i},\lambda,G^t} &= \Omega_{0,\lambda}^{k-1} \times \Omega_{j,\lambda} \times \Omega_{0,\lambda}^{t-k}, \quad \boldsymbol{i} = (k, j) \in \boldsymbol{I} := \{1, \dots, t\} \times \{1, 2, 3\},
\end{aligned}$$

(we make the reader attentive to the fact that some of the sets above are defined using complements) and the same arguments as in the $\mathrm{Sp}_r$ case give the conclusion. □

**6A2.** *Application of the large sieve.*

**Corollary 6.3.** *Let $X$, $E$, $\mathcal{O}$ and $\Lambda$ be as in Theorem 5.14. For every $\lambda \in \mathcal{O}$, let $\widehat{\mathcal{F}}_\lambda$ be a rank $r$ lisse sheaf of free $\mathcal{O}_\lambda$-modules on $X$, corresponding to a representation $\hat{\rho}_\lambda : \pi_1(X, \bar{\eta}) \to \mathrm{GL}_r(\mathcal{O}_\lambda)$. We assume (i) or (ii) of Theorem 5.14, and hypothesis (a), (b) or (c) of Corollary 5.16, hold for $\hat{\rho}_\lambda$. For $\boldsymbol{x} \in X(\mathbb{F}_q)^t$, let*

$$P_\lambda(\boldsymbol{x}) := \prod_{i=1}^t P_\lambda(x_i), \qquad P_\lambda(x_i) = \det(1 - T\rho_\lambda(\mathrm{Frob}_{x_i,q})).$$

*Then, for every $t \geq 1$ and every finite field $\mathbb{F}_q$ of characteristic $p$, we have*

$$\begin{aligned}
&\frac{|\{\boldsymbol{x} \in X(\mathbb{F}_q)^t : P_\lambda(\boldsymbol{x}) \in \mathcal{O}_\lambda[T] \text{ has nonmaximal Galois group for all } \lambda \in \Lambda\}|}{|X(\mathbb{F}_q)|^t} \qquad (25)\\
&\ll \frac{t^2}{(1-\delta_{r,t})\delta_\Lambda} \begin{cases} \sup_{\lambda\in\Lambda} \mathrm{cond}(\mathcal{F}_\lambda)^t \dfrac{\log q}{q^{1/(tE_G)}} & \text{under (a)},\\ t(B_1^2 d r^{\delta_{G=\mathrm{SL}_r}})^t(\log(B_2)M_G + \dim G)\dfrac{\log\log q}{\log q} & \text{under (b)},\\ \left(r^{\delta_{G=\mathrm{SL}_r}+1}C(X, \rho_{\lambda_0})\right)^t \dfrac{\log q}{q^{1/(2t(\dim G+1))}} & \text{under (c)}, \end{cases}
\end{aligned}$$

*with an absolute implied constant.*

*Proof.* By Proposition 6.2, the density on the left-hand side is less than or equal to

$$\sum_{\boldsymbol{i}\in\boldsymbol{I}} \frac{|\{\boldsymbol{x}\in X(\mathbb{F}_q)^t : (\rho_\lambda(\mathrm{Frob}_{x_i,q}))_{1\le i\le t}\in \boldsymbol{\Omega}_{\boldsymbol{i},\lambda,G^t} \text{ for all } \lambda\in\Lambda\}|}{|X(\mathbb{F}_q)|^t},$$

and it suffices to apply Corollary 5.16 to each summand. □

**6B.** ***Girstmair's method.*** Below, we recall the following forms of Girstmair's results [1982; 1999], as exposed in [Kowalski 2008b] (with some changes in the symmetric case).

**Definition 6.4.** For a set $M$ of complex numbers, let

$$\mathrm{Rel}_m(M)=\left\{(n_\alpha)\in\mathbb{Z}^M : \prod_{\alpha\in M}\alpha^{n_\alpha}=1\right\}.$$

**Proposition 6.5.** *Let $E$ be a number field, $t\ge 1$ an integer, and for $1\le i\le t$, let $P_i\in E[X]$ be a polynomial with splitting field $K_i$, set of roots $M_i\subset K_i$, and Galois group $G_i:=\mathrm{Gal}(K_i/E)$. We assume that the fields $K_i$ are linearly disjoint, and we let $M=\bigcup_{i=1}^t M_i$, $K=K_1\cdots K_t$. Then $\mathrm{Rel}_m(M)\otimes\mathbb{Q}=\bigoplus_{i=1}^t \mathrm{Rel}_m(M_i)\otimes\mathbb{Q}$. Moreover*:

(1) (*$W$ case*) *Assume that $G_i\cong W_r$ for some $r\ge 4$ even, acting by permutation on $M_i$. If $|\alpha|=1$ for every $\alpha\in M_i$, then*

$$\mathrm{Rel}_m(M_i)\otimes\mathbb{Q}=\{(n_\alpha)\in\mathbb{Q}^{M_i} : n_\alpha=n_{\bar\alpha}\}.$$

(2) (*$\mathfrak{S}$ case*) *Assume that $G_i\cong\mathfrak{S}_r$ for some $r\ge 2$, acting by permutation on $M_i$. Then $\mathrm{Rel}_m(M_i)\otimes\mathbb{Q}$ is either*:

(a) *if $r=2$:* $0$, $\mathbb{Q}\mathbf{1}$, *or* $\mathbb{Q}(-1,1)$.

(b) *if $r\ge 3$:* $0$ *or* $\mathbb{Q}\mathbf{1}$.

*Proof.* The $W$ case is [Kowalski 2008b, Proposition 2.4, (2.5)]. However, $\mathbb{Q}$ in the paragraph after the second display of [Kowalski 2008b, p. 13] should probably be replaced by $E$, and the contradiction comes from the fact that the splitting field of $K/E$ would be a 2-group.

For the $\mathfrak{S}$ case, note that the permutation representation $F(M_i)$ of $\mathfrak{S}_r$ decomposes as the sum of two irreducible representations

$$F(M_i)=\mathbb{Q}\mathbf{1}\oplus G(M_i),\qquad \text{where } G(M_i)=\left\{(n_\alpha)\in\mathbb{Q}^{M_i} : \sum_{\alpha\in M_i} n_\alpha=0\right\}.$$

If $G(M_i)$ is contained in the subrepresentation $\mathrm{Rel}_m(M_i)\otimes\mathbb{Q}$ of $F(M_i)$, then there exists $m\ge 1$ such that $(\alpha_j/\alpha_1)^m=1$ for $1\le j\le r$, if $M_i=\{\alpha_1,\dots,\alpha_r\}$, so that $\alpha_1^{nm}=N_{M_i/E}(\alpha_1)^m\in E$. Hence, $K_i/E$ is a Kummer extension and $\mathrm{Gal}(K_i/E)$ is abelian, which implies that $r=|M_i|=2$. If $r=2$, note that $\mathrm{Rel}_m(M_i)\otimes\mathbb{Q}=\mathbb{Q}^2$ would imply that $\mathrm{Rel}_m(M_i)=\mathbb{Z}^2$, which is a contradiction. □

### 6C. *Conclusion.*

**Corollary 6.6.** *Under the hypotheses of Corollary 6.3, assume moreover that* $(\mathcal{F}_\lambda)_{\lambda\in\Lambda}$ *forms a* ***compatible system****, i.e., that for all* $x \in X(\mathbb{F}_q)$, $P_\lambda(x) = P(x) \in E[T]$ *does not depend on* $\lambda$. *For every* $\boldsymbol{x} \in X(\mathbb{F}_q)^t$ *and* $1 \le i \le t$, *let* $M(x_i) \subset \mathbb{C}$ *be the set of zeros of* $\det(1 - T\rho_\lambda(\mathrm{Frob}_{x_i,q}))$, *so that the set of zeros of* $P_\lambda(\boldsymbol{x})$ *is* $\bigcup_{i=1}^t M(x_i)$. *Then, for all but at most a proportion* (25) *of* $\boldsymbol{x} \in X(\mathbb{F}_q)^t$, *we have*

$$\mathrm{Rel}_m(M(\boldsymbol{x})) = \begin{cases} \bigotimes_{i=1}^t \mathbb{Z}\mathbf{1} & \text{for } G = \mathrm{SL}_r \ (r \ge 2), \\ \bigotimes_{i=1}^t \{(n_\alpha) \in \mathbb{Z}^{M(x_i)} : n_\alpha = n_{\bar\alpha}\} & \text{for } G = \mathrm{Sp}_r \ (r \ge 4 \text{ even}). \end{cases}$$

*In other words, the only multiplicative relations among the roots are the trivial ones. If we write the roots of* $P(x_i)$ *as*

$$\begin{cases} e(\theta_j(x_i)) & \text{for } G = \mathrm{SL}_r \quad (1 \le j \le r), \\ e(\pm\theta_j(x_i)) & \text{for } G = \mathrm{Sp}_r \quad (1 \le j \le r/2), \end{cases}$$

*then the angles*

$$\begin{cases} 1,\ \theta_j(x_i) & \text{for } G = \mathrm{SL}_r \quad (1 \le i \le t,\ 1 \le j \le r-1), \\ 1,\ \theta_j(x_i) & \text{for } G = \mathrm{Sp}_r \quad (1 \le i \le t,\ 1 \le j \le r/2) \end{cases}$$

*are* $\mathbb{Q}$*-linearly independent for all but at most a proportion* (25) *of* $\boldsymbol{x} \in X(\mathbb{F}_q)^t$.

*Proof.* By the compatibility assumption and Corollary 6.3, $P_\lambda(\boldsymbol{x})$ has maximal Galois group $\mathfrak{S}_r^t$ or $W_r^t$ for all but at most a proportion (25) elements $\boldsymbol{x} \in X(\mathbb{F}_q)^t$. Let us assume this maximality condition holds, in which case the hypotheses of Proposition 6.5 hold. Since the product of the zeros of $P_\lambda(x_i)$ is equal to 1, we have $\mathbb{Z}\mathbf{1} \subset \mathrm{Rel}_m(M(x_i))$ for all $x_i \in X(\mathbb{F}_q)$. By Proposition 6.5 and the fact that $\mathrm{Rel}_m(M(x_i))$ is a lattice, this implies that $\mathrm{Rel}_m(M(\boldsymbol{x}))$ is as given in the statement. □

## 7. Proof of the generic linear independence theorems

In this section, we finally prove Theorems 1.1, 1.11 and 1.20, by applying Corollary 6.6. That basically means checking that assumptions (i) or (ii) of Theorem 5.14 (on monodromy groups) apply, as well as hypothesis (a) or (c) of Corollary 5.16.

### 7A. *Proof of Theorem 1.1 (exponential sums).*

Theorem 5.14(i) holds by Theorems 5.8 and 5.10 for Kloosterman sums and Birch sums respectively, with the set of valuations $\Lambda_{r,p}$, $\Lambda_p$ given therein. For Kloosterman sums, the dependency with respect to $p$ can be removed by Remark 5.9.

Since the sheaves are on curves, Corollary 5.16(a) holds. By [Katz 1988, Theorem 4.1.1(3,4)], $\mathrm{cond}(\mathrm{Kl}_{r,\lambda})$ is bounded by a constant depending only on $r$ (and not on $p$), and the same holds true for Birch sheaves by the bounds on Swan conductors and ramification points in [Katz 1990, Chapter 7]. □

### 7B. *Proof of Theorem 1.11 (super-even primitive characters).*

Theorem 5.14(i) applies by Theorem 5.12, with the set of valuations $\Lambda_{k,p}$ given by the latter.

If $p > k$, we see (as in [Katz 2017, Lemma 5.2]) that $\mathbb{W}_{2\kappa,\,\mathrm{odd}} = \prod_{1\le a\le 2\kappa,\,a\ \mathrm{odd}} W_1$ is the space of odd polynomials of degree $\le 2\kappa - 1$ and $\mathrm{Prim}_{2\kappa,\,\mathrm{odd}}$ the subspace of those polynomials with degree exactly $2\kappa - 1$. One can then apply Corollary 5.16(c), which gives the theorem.

To obtain the weaker error (but with explicit base of $t$) in Remark 1.13, one applies Corollary 5.16(b) instead, using the bounds for Betti numbers in [Katz 2017, Lemma 5.2], giving $B_1 = 3(2\kappa+1)^{2\kappa}$, $B_2 = 2\kappa + 1$. □

**7C.** ***Proof of Theorem 1.20 (even primitive characters).*** In this case, hypothesis (ii) of Theorem 5.14 (projective monodromy groups) applies by Theorem 5.12.

If $p > m$, then as in [Katz 2017], we see that $\mathbb{W}_m = \prod_{1\le a\le m} W_1$ is the space of polynomials of degree $\le m$ with constant term 1 and $\mathrm{Prim}_m$ is the subspace of those polynomials with degree exactly $m$. One can then apply Corollary 5.16(c), which gives the theorem.

To obtain the weaker error (but with explicit base of $t$) in Remark 1.26, one applies Corollary 5.16(c) instead, proceeding from [Katz 2013b] as in [Katz 2017, Lemma 5.2] to bound the Betti numbers. Let us indeed show that Hypothesis (b) of Corollary 5.16 holds with $B_1 = 3(m+1)^{m+1}$ and $B_2 = m+1$. Let $M \ge 1$ be an integer. With coordinates $(t_1, \dots, t_M, f)$ on $\mathbb{A}^M \times \mathrm{Prim}_m$,

$$H_c^i(\mathrm{Prim}_m, \mathcal{L}_{\mathrm{univ}}^{\otimes M}) = H_c^{i+M}(\mathbb{A}^M \times \mathrm{Prim}_m, \mathcal{L}_{\psi(f(t_1)+\cdots+f(t_M))}).$$

Note that $\mathbb{A}^M \times \mathrm{Prim}_m$ is defined in $\mathbb{A}^{M+1+m}$ (an additional coordinate is needed for the condition that $a_m \ne 0$) and $f(t_1)+\cdots+f(t_M)$ is a polynomial in $t_i, a_i$ of degree $m+1$. By [Katz 2001, Theorem 12] (with $(\delta, N, r, d, s, e_j) = (m+1, M+1+m, 1, 2, 0, 0)$), we have

$$\sigma_c(\mathrm{Prim}_m, \mathcal{L}_{\mathrm{univ}}^{\otimes M}) \le 3(1+\max(m+1,3))^{M+m+1} = 3(m+1)^{M+m+1}.$$ □

## Acknowledgements

The author thanks Lucile Devin, Michele Fornea, Javier Fresán, Florent Jouve and Will Sawin for helpful discussions and comments. Will Sawin in particular provided a better way to bound the sums of Betti numbers in the large sieve, leading to stronger results; the idea and proof of Theorem 5.14(c) are due to him. We thank the organizers of the 2019 Shaoul fund IAS Function field arithmetic workshop in Tel-Aviv for providing the opportunity for some of these exchanges. We are grateful to the anonymous referees who provided helpful and detailed comments to improve the manuscript. The author was partially supported by Koukoulopoulos' Discovery Grant 435272-2013 of the Natural Sciences and Engineering Research Council of Canada, and by Radziwiłł's NSERC DG grant and the CRC program.

## References

[Ahmadi and Shparlinski 2010] O. Ahmadi and I. E. Shparlinski, "On the distribution of the number of points on algebraic curves in extensions of finite fields", *Math. Res. Lett.* **17**:4 (2010), 689–699. MR Zbl

[Baker and Wüstholz 1993] A. Baker and G. Wüstholz, "Logarithmic forms and group varieties", *J. Reine Angew. Math.* **442** (1993), 19–62. MR Zbl

[Billingsley 1986] P. Billingsley, *Probability and measure*, 2nd ed., Wiley, New York, 1986. MR Zbl

[Bombieri and Katz 2010] E. Bombieri and N. M. Katz, "A note on lower bounds for Frobenius traces", *Enseign. Math.* (2) **56**:3-4 (2010), 203–227. MR Zbl

[Bourbaki 2005] N. Bourbaki, *Lie groups and Lie algebras: Chapters 7–9*, Springer, 2005. MR Zbl

[Brauer 1964] R. Brauer, "A note on theorems of Burnside and Blichfeldt", *Proc. Amer. Math. Soc.* **15** (1964), 31–34. MR Zbl

[Bryant and Kovács 1972] R. M. Bryant and L. G. Kovács, "Tensor products of representations of finite groups", *Bull. London Math. Soc.* **4** (1972), 133–135. MR Zbl

[Cha 2008] B. Cha, "Chebyshev's bias in function fields", *Compos. Math.* **144**:6 (2008), 1351–1374. MR Zbl

[Cha and Kim 2010] B. Cha and S. Kim, "Biases in the prime number race of function fields", *J. Number Theory* **130**:4 (2010), 1048–1055. MR Zbl

[Cha et al. 2016] B. Cha, D. Fiorilli, and F. Jouve, "Prime number races for elliptic curves over function fields", *Ann. Sci. Éc. Norm. Supér.* (4) **49**:5 (2016), 1239–1277. MR Zbl

[Cha et al. 2017] B. Cha, D. Fiorilli, and F. Jouve, "Independence of the zeros of elliptic curve $L$-functions over function fields", *Int. Math. Res. Not.* **2017**:9 (2017), 2614–2661. MR Zbl

[Curtis and Reiner 1962] C. W. Curtis and I. Reiner, *Representation theory of finite groups and associative algebras*, Pure and Applied Mathematics **XI**, Interscience, New York, 1962. MR Zbl

[Devin 2019] L. Devin, "Chebyshev's bias for analytic $L$-functions", *Math. Proc. Cambridge Philos. Soc.* (online publication March 2019).

[Devin and Meng 2018] L. Devin and X. Meng, "Chebyshev's bias for products of irreducible polynomials", preprint, 2018. arXiv

[Evertse 1984] J.-H. Evertse, "On sums of $S$-units and linear recurrences", *Compositio Math.* **53**:2 (1984), 225–244. MR Zbl

[Fouvry et al. 2015] E. Fouvry, E. Kowalski, and P. Michel, "A study in sums of products", *Philos. Trans. Roy. Soc. A* **373**:2040 (2015), art. id. 20140309, 26 pp. MR Zbl

[Fu 2011] L. Fu, *Etale cohomology theory*, Nankai Tracts in Mathematics **13**, World Scientific, Hackensack, NJ, 2011. MR Zbl

[Gallagher 1973] P. X. Gallagher, "The large sieve and probabilistic Galois theory", pp. 91–101 in *Analytic number theory* (St. Louis, MO, 1972), edited by H. G. Diamond, Proc. Sympos. Pure Math. **XXIV**, Amer. Math. Soc., Providence, RI, 1973. MR Zbl

[Girstmair 1982] K. Girstmair, "Linear dependence of zeros of polynomials and construction of primitive elements", *Manuscripta Math.* **39**:1 (1982), 81–97. MR Zbl

[Girstmair 1999] K. Girstmair, "Linear relations between roots of polynomials", *Acta Arith.* **89**:1 (1999), 53–96. MR Zbl

[Gluck 1993] D. Gluck, "Character value estimates for nonsemisimple elements", *J. Algebra* **155**:1 (1993), 221–237. MR Zbl

[Gouillon 2006] N. Gouillon, "Explicit lower bounds for linear forms in two logarithms", *J. Théor. Nombres Bordeaux* **18**:1 (2006), 125–146. MR Zbl

[Hall 2008] C. Hall, "Big symplectic or orthogonal monodromy modulo $l$", *Duke Math. J.* **141**:1 (2008), 179–203. MR Zbl

[Humphreys 2006] J. E. Humphreys, *Modular representations of finite groups of Lie type*, London Mathematical Society Lecture Note Series **326**, Cambridge University Press, 2006. MR Zbl

[Illusie 1981] L. Illusie, "Théorie de Brauer et caractéristique d'Euler–Poincaré (d'après P. Deligne)", pp. 161–172 in *The Euler–Poincaré characteristic*, Astérisque **82**, Soc. Math. France, Paris, 1981. MR Zbl

[Jouve et al. 2013] F. Jouve, E. Kowalski, and D. Zywina, "Splitting fields of characteristic polynomials of random elements in arithmetic groups", *Israel J. Math.* **193**:1 (2013), 263–307. MR Zbl

[Katz 1987] N. M. Katz, "On the monodromy groups attached to certain families of exponential sums", *Duke Math. J.* **54**:1 (1987), 41–56. MR Zbl

[Katz 1988] N. M. Katz, *Gauss sums, Kloosterman sums, and monodromy groups*, Annals of Mathematics Studies **116**, Princeton University Press, 1988. MR Zbl

[Katz 1990] N. M. Katz, *Exponential sums and differential equations*, Annals of Mathematics Studies **124**, Princeton University Press, 1990. MR Zbl

[Katz 2001] N. M. Katz, "Sums of Betti numbers in arbitrary characteristic", *Finite Fields Appl.* **7**:1 (2001), 29–44. MR Zbl

[Katz 2002] N. M. Katz, *Twisted L-functions and monodromy*, Annals of Mathematics Studies **150**, Princeton University Press, 2002. MR Zbl

[Katz 2005] N. M. Katz, *Moments, monodromy, and perversity: a Diophantine perspective*, Annals of Mathematics Studies **159**, Princeton University Press, 2005. MR Zbl

[Katz 2012a] N. M. Katz, *Convolution and equidistribution: Sato-Tate theorems for finite-field Mellin transforms*, Annals of Mathematics Studies **180**, Princeton University Press, 2012. MR Zbl

[Katz 2012b] N. M. Katz, "Report on the irreducibility of $L$-functions", pp. 321–353 in *Number theory, analysis and geometry*, edited by D. Goldfeld et al., Springer, 2012. MR Zbl

[Katz 2013a] N. M. Katz, "On a question of Keating and Rudnick about primitive Dirichlet characters with squarefree conductor", *Int. Math. Res. Not.* **2013**:14 (2013), 3221–3249. MR Zbl

[Katz 2013b] N. M. Katz, "Witt vectors and a question of Keating and Rudnick", *Int. Math. Res. Not.* **2013**:16 (2013), 3613–3638. MR Zbl

[Katz 2017] N. M. Katz, "Witt vectors and a question of Rudnick and Waxman", *Int. Math. Res. Not.* **2017**:11 (2017), 3377–3412. MR Zbl

[Katz and Sarnak 1999] N. M. Katz and P. Sarnak, *Random matrices, Frobenius eigenvalues, and monodromy*, American Mathematical Society Colloquium Publications **45**, Amer. Math. Soc., Providence, RI, 1999. MR Zbl

[Keating and Rudnick 2014] J. P. Keating and Z. Rudnick, "The variance of the number of prime polynomials in short intervals and in residue classes", *Int. Math. Res. Not.* **2014**:1 (2014), 259–288. MR Zbl

[Kerz and Schmidt 2010] M. Kerz and A. Schmidt, "On different notions of tameness in arithmetic geometry", *Math. Ann.* **346**:3 (2010), 641–668. MR Zbl

[Kowalski 2006] E. Kowalski, "The large sieve, monodromy and zeta functions of curves", *J. Reine Angew. Math.* **601** (2006), 29–69. MR Zbl

[Kowalski 2008a] E. Kowalski, *The large sieve and its applications: arithmetic geometry, random walks and discrete groups*, Cambridge Tracts in Mathematics **175**, Cambridge University Press, 2008. MR Zbl

[Kowalski 2008b] E. Kowalski, "The large sieve, monodromy, and zeta functions of algebraic curves, II: Independence of the zeros", *Int. Math. Res. Not.* **2008** (2008), art. id. rnn091, 57 pp. MR Zbl

[Larsen 1995] M. Larsen, "Maximality of Galois actions for compatible systems", *Duke Math. J.* **80**:3 (1995), 601–630. MR Zbl

[Larsen and Pink 1992] M. Larsen and R. Pink, "On $l$-independence of algebraic monodromy groups in compatible systems of representations", *Invent. Math.* **107**:3 (1992), 603–636. MR Zbl

[Li 2018] W. Li, "Vanishing of hyperelliptic L-functions at the central point", *J. Number Theory* **191** (2018), 85–103. MR Zbl

[Livné 1987] R. Livné, "The average distribution of cubic exponential sums", *J. Reine Angew. Math.* **375/376** (1987), 362–379. MR Zbl

[Malle and Testerman 2011] G. Malle and D. Testerman, *Linear algebraic groups and finite groups of Lie type*, Cambridge Studies in Advanced Mathematics **133**, Cambridge University Press, 2011. MR Zbl

[Martin and Ng 2017] G. Martin and N. Ng, "Inclusive prime number races", preprint, 2017. arXiv

[Matthews et al. 1984] C. R. Matthews, L. N. Vaserstein, and B. Weisfeiler, "Congruence properties of Zariski-dense subgroups, I", *Proc. London Math. Soc.* (3) **48**:3 (1984), 514–532. MR Zbl

[Narkiewicz 2004] W. a. a. Narkiewicz, *Elementary and analytic theory of algebraic numbers*, 3rd ed., Springer, 2004. MR Zbl

[Perret-Gentil 2017] C. Perret-Gentil, "Gaussian distribution of short sums of trace functions over finite fields", *Math. Proc. Cambridge Philos. Soc.* **163**:3 (2017), 385–422. MR Zbl

[Perret-Gentil 2018a] C. Perret-Gentil, "Exponential sums over finite fields and the large sieve", *Int. Math. Res. Not.* (online publication August 2018), art. id. rny202, 36 pp.

[Perret-Gentil 2018b] C. Perret-Gentil, "Integral monodromy groups of Kloosterman sheaves", *Mathematika* **64**:3 (2018), 652–678. MR Zbl

[Pink 2000] R. Pink, "Strong approximation for Zariski dense subgroups over arbitrary global fields", *Comment. Math. Helv.* **75**:4 (2000), 608–643. MR Zbl

[van der Poorten and Schlickewei 1991] A. J. van der Poorten and H. P. Schlickewei, "Additive relations in fields", *J. Austral. Math. Soc. Ser. A* **51**:1 (1991), 154–170. MR Zbl

[Rains 1997] E. M. Rains, "High powers of random elements of compact Lie groups", *Probab. Theory Related Fields* **107**:2 (1997), 219–241. MR Zbl

[Rosen 2002] M. Rosen, *Number theory in function fields*, Graduate Texts in Mathematics **210**, Springer, 2002. MR Zbl

[Rubinstein and Sarnak 1994] M. Rubinstein and P. Sarnak, "Chebyshev's bias", *Experiment. Math.* **3**:3 (1994), 173–197. MR Zbl

[Rudnick and Waxman 2019] Z. Rudnick and E. Waxman, "Angles of Gaussian primes", *Israel J. Math.* **232**:1 (2019), 159–199. MR Zbl

[SGA $4\frac{1}{2}$ 1977] P. Deligne, *Cohomologie étale* (Séminaire de Géométrie Algébrique du Bois Marie), Lecture Notes in Math. **569**, Springer, 1977. MR Zbl

[Speyer 2011] D. E. Speyer, "Faithful representations and tensor powers", answer on MathOverflow, 2011, available at https://mathoverflow.net/q/63043.

[Stein 1993] E. M. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, Princeton Mathematical Series **43**, Princeton University Press, 1993. MR Zbl

[Steinberg 1962] R. Steinberg, "Complete sets of representations of algebras", *Proc. Amer. Math. Soc.* **13** (1962), 746–747. MR Zbl

[Weisfeiler 1984] B. Weisfeiler, "Strong approximation for Zariski-dense subgroups of semisimple algebraic groups", *Ann. of Math.* (2) **120**:2 (1984), 271–315. MR Zbl

Communicated by Melanie Matchett Wood
Received 2019-05-09 Revised 2019-09-14 Accepted 2019-12-16

corentin.perretgentil@gmail.com *Zürich, Switzerland*

## Guidelines for Authors

Authors may submit manuscripts in PDF format on-line at the Submission page at the ANT website.

**Originality**. Submission of a manuscript acknowledges that the manuscript is original and and is not, in whole or in part, published or under consideration for publication elsewhere. It is understood also that the manuscript will not be submitted elsewhere while under consideration for publication in this journal.

**Language**. Articles in *ANT* are usually in English, but articles written in other languages are welcome.

**Length** There is no a priori limit on the length of an *ANT* article, but *ANT* considers long articles only if the significance-to-length ratio is appropriate. Very long manuscripts might be more suitable elsewhere as a memoir instead of a journal article.

**Required items**. A brief abstract of about 150 words or less must be included. It should be self-contained and not make any reference to the bibliography. If the article is not in English, two versions of the abstract must be included, one in the language of the article and one in English. Also required are keywords and subject classifications for the article, and, for each author, postal address, affiliation (if appropriate), and email address.

**Format**. Authors are encouraged to use LaTeX but submissions in other varieties of TeX, and exceptionally in other formats, are acceptable. Initial uploads should be in PDF format; after the refereeing process we will ask you to submit all source material.

**References**. Bibliographical references should be complete, including article titles and page ranges. All references in the bibliography should be cited in the text. The use of BibTeX is preferred but not required. Tags will be converted to the house format, however, for submission you may use the format of your choice. Links will be provided to all literature with known web locations and authors are encouraged to provide their own links in addition to those supplied in the editorial process.

**Figures**. Figures must be of publication quality. After acceptance, you will need to submit the original source files in vector graphics format for all diagrams in your manuscript: vector EPS or vector PDF files are the most useful.

Most drawing and graphing packages (Mathematica, Adobe Illustrator, Corel Draw, MATLAB, etc.) allow the user to save files in one of these formats. Make sure that what you are saving is vector graphics and not a bitmap. If you need help, please write to graphics@msp.org with details about how your graphics were generated.

**White space**. Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

**Proofs**. Page proofs will be made available to authors (or to the designated corresponding author) at a Web site in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

Algebra & Number Theory

Volume 14 No. 5 2020



1937-0652(2020)14:5;1-X