Algebra & Number Theory

Volume 13
2019
No. 6

msp

Algebra & Number Theory

msp.org/ant

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

The subscription price for 2019 is US $385/year for the electronic version, and $590/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/

ALGEBRA AND NUMBER THEORY 13:6 (2019)
dx.doi.org/10.2140/ant.2019.13.1243

Positivity functions for curves on algebraic varieties

Brian Lehmann and Jian Xiao

This is the second part of our work on Zariski decomposition structures, where we compare two different volume type functions for curve classes. The first function is the polar transform of the volume for divisor classes. The second function captures the asymptotic geometry of curves analogously to the volume function for divisors. We prove that the two functions coincide, generalizing Zariski's classical result for surfaces to all varieties. Our result confirms the log concavity conjecture of the first named author for weighted mobility of curve classes in an unexpected way, via Legendre–Fenchel type transforms. During the course of the proof, we obtain a refined structure theorem for the movable cone of curves.

1. Introduction

Let X be a smooth complex projective variety of dimension n. The Riemann–Roch problem asks whether one can determine the dimension of the space of sections of a holomorphic line bundle L on X. An important subtlety of this problem is that the answer is not determined by purely topological data — line bundles which share the same Chern class need not have isomorphic spaces of sections. In general, the problem only has a satisfactory answer for sufficiently ample line bundles, which exhibit a close relationship between geometry, cohomology, and intersection theory.

Over the past forty years, mathematicians have realized that one obtains a much richer theory by studying the asymptotic behavior of the space of sections of mL as m increases. Indeed, by working asymptotically, we can recover for general effective line bundles some of the same interplay between sheaf cohomology and intersection theory which undergirds the theory of ample line bundles. This point

MSC2010: primary 14C25; secondary 14C20, 32J25.
Keywords: algebraic varieties, positivity of curves, mobility of cycles, volume-type function, Zariski decomposition.

of view leads to many important positivity invariants for line bundles and linear systems. Perhaps the most important asymptotic invariant of a line bundle L is its volume,[1] defined as

$$\mathrm{vol}(L) := \limsup_{m\to\infty} \frac{\dim H^0(X, mL)}{m^n/n!}.$$

When X is a surface, the volume of L can be calculated using intersection theory. The key construction is the Zariski decomposition [1962], which splits L into a "positive" part and a "rigid" part. In higher dimensions as well, there is a close relationship between the asymptotic geometry of divisors and intersection-theoretic positivity via volume-type functions.

Recently there has been interest in extending the theory of positivity to subvarieties of arbitrary codimension (see e.g., [Debarre et al. 2011; Lehmann 2016; Fulger and Lehmann 2017a; 2017b]). By analogy, one would like to study the asymptotic geometry of cycles and its relationship with numerical measures of positivity. In this paper we develop such a theory for curves: we show that the asymptotic enumerative geometry of curve classes is controlled by intersection-theoretic invariants.

Our comparison relies upon several natural volume-type functions for curve classes. The first function involves the numerical positivity of a curve class.

Definition 1.1 [Xiao 2017, Definition 1.1]. Let X be a projective variety of dimension n and let $\alpha \in \overline{\mathrm{Eff}}_1(X)$ be a pseudoeffective curve class. Then the volume of α is defined to be

$$\widehat{\mathrm{vol}}(\alpha) = \inf_{A \text{ big and nef divisor class}} \left(\frac{A \cdot \alpha}{\mathrm{vol}(A)^{1/n}} \right)^{n/n-1}.$$

When α is a curve class that is not pseudoeffective, we set $\widehat{\mathrm{vol}}(\alpha) = 0$.

This is a polar transformation of the volume function on the ample cone of divisors. The definition is inspired by the realization that the volume of a divisor has a similar intersection-theoretic description against curves as in [Xiao 2017, Theorem 2.1]. It fits into a much broader picture relating positivity of divisors and curves via cone duality; see [Lehmann and Xiao 2016].

The second function captures the asymptotic geometry of curves. Recall that general points impose independent codimension 1 conditions on divisors in a linear series. Thus for a divisor L, one can interpret $\dim \mathbb{P}(H^0(X, L))$ as a measurement of how many general points are contained in sections of L. Using this interpretation, we define the mobility function for curves in an analogous way.

Definition 1.2 [Lehmann 2016, Definition 1.1]. Let X be a projective variety of dimension n and let $\alpha \in N_1(X)$ be a curve class with integer coefficients. The mobility of α is defined to be

$$\mathrm{mob}(\alpha) := \limsup_{m\to\infty} \frac{\max\{b \in \mathbb{Z}_{\geq 0} \mid \text{any } b \text{ general points are contained in an effective curve of class } m\alpha\}}{m^{n/(n-1)}/n!}.$$

There is a closely related function known as the weighted mobility which counts singular points of the curve with a "higher weight". We first recall the definition of the weighted mobility count for a class

[1] For a nonbig line bundle, the higher asymptotic cohomological functions carry more significant information [Küronya 2006].

$\alpha \in N_1(X)$ with integer coefficients (see [Lehmann 2016, Definition 8.6]):

$$\mathrm{wmc}(\alpha) = \sup_{\mu} \max \left\{ b \in \mathbb{Z}_{\geq 0} \;\middle|\; \begin{array}{l} \text{there is an effective cycle of class } \mu\alpha \text{ through any } b \\ \text{points of } X \text{ with multiplicity at least } \mu \text{ at each point} \end{array} \right\}.$$

The supremum is shown to exist in [Lehmann 2016] — it is then clear that the supremum is achieved by some positive integer μ. We define the weighted mobility to be

$$\mathrm{wmob}(\alpha) = \limsup_{m\to\infty} \frac{\mathrm{wmc}(m\alpha)}{m^{n/n-1}}.$$

While the definition is slightly more complicated, the weighted mobility is easier to compute due to its close relationship with Seshadri constants. Lehmann [2016] showed that both the mobility and weighted mobility extend to continuous homogeneous functions on all of $N_1(X)$.

Main result. Our main theorem compares these functions. It continues a project begun in [Xiao 2017] (see especially Conjecture 3.1 and Theorem 3.2 there).

Theorem 1.3 (see Theorem 6.1). *Let X be a smooth projective variety of dimension n and let $\alpha \in \overline{\mathrm{Eff}}_1(X)$ be a pseudoeffective curve class. Then*:

(1) $\widehat{\mathrm{vol}}(\alpha) = \mathrm{wmob}(\alpha)$.

(2) $\widehat{\mathrm{vol}}(\alpha) \leq \mathrm{mob}(\alpha) \leq n!\,\widehat{\mathrm{vol}}(\alpha)$.

(3) *Assume Conjecture 1.4 below. Then* $\mathrm{mob}(\alpha) = \widehat{\mathrm{vol}}(\alpha)$.

This result is surprising: it suggests that the mobility count of *any* curve class is optimized by complete intersection curves; see the end of Section 2 (page 1256). Just as for curves on algebraic surfaces, the key to this result is the Zariski decomposition for curves on varieties of arbitrary dimension as constructed in [Lehmann and Xiao 2016; Fulger and Lehmann 2017b]. Part (3) of the theorem relies on the following conjectural description of the mobility of a complete intersection class:

Conjecture 1.4 [Lehmann 2016, Question 6.1]. *Let X be a smooth projective variety of dimension n and let A be an ample divisor on X. Then*

$$\mathrm{mob}(A^{n-1}) = A^n.$$

Example 1.5. Let α denote the class of a line on $\mathbb{P}^3$. The mobility count of α is determined by the following enumerative question: what is the minimal degree of a curve through b general points of $\mathbb{P}^3$? The answer is unknown, even in an asymptotic sense.

Perrin [1987] conjectured that the "optimal" curves (which maximize the number of points relative to their degree to the $\frac{3}{2}$) are complete intersections of two divisors of the same degree. Theorem 1.3 supports a vast generalization of Perrin's conjecture to all big curve classes on all smooth projective varieties.

Strict log concavity of the volume function for divisors. An important ingredient in the proof of Theorem 1.3 is the study of the volume function for divisors from the perspective of convexity theory. Since such results are of interest in their own right, we summarize the highlights below.

The first step is to analyze the strict log concavity of the volume function. It is well-known that the volume function for divisor classes is log concave (see e.g., [Lazarsfeld 2004, Theorem 11.4.9; Boucksom 2002b]). We show that it is strictly log concave on the big and movable cone of divisors (but on no larger cone), extending [Boucksom et al. 2009, Theorem D].

Theorem 1.6. *Let X be a smooth projective variety of dimension n. For any two big divisor classes L_1, L_2, the inequality*

$$\mathrm{vol}(L_1+L_2)^{1/n} \geq \mathrm{vol}(L_1)^{1/n} + \mathrm{vol}(L_2)^{1/n}$$

is an equality if and only if the (numerical) positive parts $P_\sigma(L_1)$, $P_\sigma(L_2)$ are proportional. Thus the function $L \mapsto \mathrm{vol}(L)$ is strictly log concave on the cone of big and movable divisors.

This result is proved in Section 3 (see Theorem 3.9). It shows that the volume function for divisors fits into the abstract convexity framework developed in [Lehmann and Xiao 2016]. A posteriori, this viewpoint motivates many of the well-known structure results for the volume function (such as the formula for the derivative, the Khovanskii–Teissier inequalities, the σ-decomposition, etc.).

Refined structure of the movable cone. The most important consequence is a refined version of a theorem of [Boucksom et al. 2013] describing the movable cone of curves. In [loc. cit.], it is proved that the movable cone $\mathrm{Mov}_1(X)$ is the closure of the cone *generated* by $(n-1)$-self positive products of big divisors. We show that $\mathrm{Mov}_1(X)$ is the *closure* of the set of $(n-1)$-self positive products of big divisors on the interior of $\mathrm{Mov}^1(X)$. (The definition of the positive product $\langle - \rangle$ is recalled in Section 2.)

Theorem 1.7. *Let X be a smooth projective variety of dimension n. The $(n-1)$-st positive product $\langle -^{n-1} \rangle$ defines a continuous bijection from the interior of the big and movable cone of divisors to the interior of* $\mathrm{Mov}_1(X)$.

In practice, Theorem 1.7 seems quite useful for working with the movable cone of curves. For example, it has an immediate corollary:

Corollary 1.8. *Let X be a projective variety of dimension n. Then the rays over classes of irreducible curves which deform to dominate X are dense in* $\mathrm{Mov}_1(X)$.

Polar transform of the volume function for divisors. Equipped with these results, we return to our discussion of positivity functions for curves.

First we review some facts about polar transforms. Let V be a real vector space of dimension n, and let V^* be its dual space. Let $\mathrm{Cvx}(V)$ be the space of lower semicontinuous convex functions on V. We denote the paring of $w^* \in V^*$ and $v \in V$ by $w^* \cdot v$. Recall that the classical Legendre–Fenchel transform

$$\mathcal{L} : \mathrm{Cvx}(V) \to \mathrm{Cvx}(V^*), \quad \mathcal{L}f(w^*) = \sup_{v \in V}\{w^* \cdot v - f(v)\}$$

is an order-reversing involution which relates the differentiability of a convex function with the strict convexity of its dual (see e.g., [Rockafellar 1970]).

When working with homogeneous functions on a cone, there is an analogue of the Legendre–Fenchel transform which plays a similar theoretical role. It is the concave homogeneous version of the well-known polar transform. Let $\mathfrak{C} \subset V$ be a proper closed convex cone of full dimension and let $\mathfrak{C}^* \subset V^*$ be its dual cone. We let $\mathrm{HConc}_s(\mathfrak{C})$ denote the collection of functions $f : \mathfrak{C} \to \mathbb{R}$, which are upper-semicontinuous, homogeneous of weight $s > 1$, strictly positive in the interior of $\mathfrak{C}$ and s-concave. The polar transform $\mathcal{H}$ associates to a function $f \in \mathrm{HConc}_s(\mathfrak{C})$ the function $\mathcal{H}f \in \mathrm{HConc}_{s/(s-1)}(\mathfrak{C}^*)$ defined as

$$\mathcal{H}f(w^*) := \inf_{v \in \mathfrak{C}^\circ} \left(\frac{w^* \cdot v}{f(v)^{1/s}} \right)^{s/(s-1)}.$$

By taking the logarithmic function of $\mathcal{H}f$, we get

$$\log \mathcal{H}f(w^*) = \frac{s}{s-1} \inf_{v \in \mathfrak{C}^\circ} \bigl(\log(w^* \cdot v) - \tfrac{1}{s} \log f(v) \bigr).$$

Thus the polar transform $\mathcal{H}$ can be considered as a variant of Legendre–Fenchel transform with a “coupling function” given by the logarithmic function. The papers [Xiao 2017; Lehmann and Xiao 2016] develop the theory of $\mathcal{H}$ in parallel with the classical Legendre–Fenchel transform $\mathcal{L}$ and demonstrate how it has fruitful applications in the positivity theory of curves.

In our geometric setting, polar duality yields two natural numerical positivity functions for curves. One is the function $\widehat{\mathrm{vol}}$ discussed above. If we instead take the polar transform of the volume on the pseudoeffective cone, then we obtain a polar function on the dual cone $\mathrm{Mov}_1(X)$.

Definition 1.9 [Xiao 2017, Definition 2.2]. Let X be a projective variety of dimension n. For any curve class $\alpha \in \mathrm{Mov}_1(X)$ define

$$\mathfrak{M}(\alpha) = \inf_{L \text{ big divisor class}} \left(\frac{L \cdot \alpha}{\mathrm{vol}(L)^{1/n}} \right)^{n/(n-1)}.$$

When α is a curve class that is not movable, we set $\mathfrak{M}(\alpha) = 0$.

While the positivity functions $\widehat{\mathrm{vol}}$, mob, wmob are conjecturally the same, $\mathfrak{M}$ exhibits quite different behavior. It is best understood as a way of making Theorem 1.7 explicit (see Lemma 3.11, Theorem 3.14 and Corollary 3.23).

Theorem 1.10. *Let X be a smooth projective variety and let α be a curve class in* $\mathrm{Mov}_1(X)$. *Then exactly one of the following alternatives holds*:

- $\alpha = \langle L^{n-1} \rangle$ *for a big movable divisor class L.*
- $\alpha \cdot M = 0$ *for a nonzero movable divisor class M.*

In the first case, we have $\mathfrak{M}(\alpha) = \mathrm{vol}(L)$ and L achieves the infimum of Definition 1.9. In the second case we have $\mathfrak{M}(\alpha) = 0$.

A curve class α of the first type lies on the boundary of $\mathrm{Mov}_1(X)$ *if and only if the corresponding big* *divisor L lies on the boundary of* $\mathrm{Mov}^1(X)$. *Thus the homeomorphism between the interiors of* $\mathrm{Mov}^1(X)$ *and* $\mathrm{Mov}_1(X)$ *given by Theorem 1.7 extends to a homeomorphism from all big movable divisor classes to curve classes with* $\mathfrak{M} > 0$.

Conceptually, the function $\mathfrak{M}$ allows us to assign a movable divisor to a movable curve class by "taking an $(n-1)$-th root". For toric varieties, this coheres with a classical construction of Minkowski which assigns a polytope to a positive Minkowski weight.

Theorem 1.3 relies upon the following comparison between the two polar transforms $\widehat{\mathrm{vol}}$ and $\mathfrak{M}$ (see Section 5). Recall that the complete intersection cone $\mathrm{CI}_1(X)$ is the closure of the set of curve classes of the form A^{n-1} for an ample divisor class A. The set $\mathrm{CI}_1(X)$ is a closed cone but may fail to be convex (see [Lehmann and Xiao 2016]).

Theorem 1.11. *Let X be a smooth projective variety and let α be a big curve class in* $\mathrm{Mov}_1(X)$. *Then the following conditions are equivalent*:

- $\alpha \in \mathrm{CI}_1(X)$.
- $\widehat{\mathrm{vol}}(\alpha) = \mathfrak{M}(\alpha)$.
- $\widehat{\mathrm{vol}}(\alpha) = \widehat{\mathrm{vol}}(\phi^*\alpha)$ *for every birational morphism* $\phi : Y \to X$.

While not strictly necessary for our main result, we also show that $\mathfrak{M}$ admits an enumerative interpretation. We define $\mathrm{mob}_{\mathrm{mov}}$ and $\mathrm{wmob}_{\mathrm{mov}}$ for curve classes analogously to mob and wmob, except that we only count contributions of families whose general member is a sum of irreducible movable curves (see paragraph after Definition 6.7 for more details).

Theorem 1.12. *Let X be a smooth projective variety of dimension n and let* $\alpha \in \mathrm{Mov}_1(X)^\circ$. *Then*:

(1) $\mathfrak{M}(\alpha) = \mathrm{wmob}_{\mathrm{mov}}(\alpha)$.

(2) *Assume Conjecture 1.4. Then* $\mathfrak{M}(\alpha) = \mathrm{mob}_{\mathrm{mov}}(\alpha)$.

Outline of the proof. We briefly outline the proof of Theorem 1.3(3), the most difficult part. As mentioned above, Zariski decompositions for positivity functions play an important role. Fix a function $f \in \{\widehat{\mathrm{vol}}, \mathrm{mob}\}$. A Zariski decomposition for a big curve class α with respect to f is an expression

$$\alpha = P + N$$

where N is pseudoeffective and P is a "positive part" satisfying $f(P) = f(\alpha)$.

The main distinction between the Zariski decompositions for mob and $\widehat{\mathrm{vol}}$ is where the positive part is required to lie. For $\widehat{\mathrm{vol}}$, the positive part $P_{\widehat{\mathrm{vol}}}$ constructed in [Lehmann and Xiao 2016] lies in the complete intersection cone $\mathrm{CI}_1(X)$. For mob, the positive part P_{mob} constructed in [Fulger and Lehmann 2017b] lies in $\mathrm{Mov}_1(X)$. In fact a stronger property is proved there: $\mathrm{mob}(P_{\mathrm{mob}}) = \mathrm{mob}(\phi^* P_{\mathrm{mob}})$ for any birational map ϕ. Using Conjecture 1.4 and a delicate comparison between mob and $\mathfrak{M}$, Theorem 1.11 allows us to conclude the stronger statement that $P_{\mathrm{mob}} \in \mathrm{CI}_1(X)$. Then the two positive parts should coincide, and one can again apply Conjecture 1.4 to deduce the equality of the two functions.

Examples.

Hyperkähler varieties. For hyperkähler varieties, positivity functions for curves admit interesting interpretations in terms of the Beauville–Bogomolov form. Let q denote the Beauville–Bogomolov quadratic form on $N^1(X)$ normalized so that $q(D)^{n/2} = D^n$ for ample D. The form induces an isomorphism $\psi : N^1(X) \to N_1(X)$. Then [Lehmann and Xiao 2016, Section 7] shows that the bijection of Theorem 1.7 can be understood using ψ:

- If D is a big movable divisor, then $\mathrm{vol}(D)^{(n-2)/n}\psi(D) = \langle D^{n-1}\rangle$. In other words, the bijection of Theorem 1.7 coincides with ψ up to a continuous rescaling factor. For a big movable divisor class D,

$$\mathfrak{M}(\psi(D)) = \mathrm{vol}(D)^{1/n-1}.$$

- In particular, ψ also induces a bijection between the big and nef cone of divisors and the complete intersection curve classes with positive $\widehat{\mathrm{vol}}$. For A big and nef we have

$$\widehat{\mathrm{vol}}(\psi(A)) = \mathrm{vol}(A)^{1/n-1}.$$

In general, the volume of a curve class is given by a Zariski decomposition projecting into the complete intersection cone. [Lehmann and Xiao 2016] furthermore shows how this decomposition is related via q-duality to the σ-decomposition of divisors.

Mori dream spaces. If X is a Mori dream space, then the movable cone of divisors admits a chamber structure defined via the ample cones on small $\mathbb{Q}$-factorial modifications. This chamber structure behaves compatibly with the σ-decomposition and the volume function for divisors.

For curves we obtain a complementary picture using the movable cone of curves. Note that $\mathrm{Mov}_1(X)$ is naturally preserved by small $\mathbb{Q}$-factorial modifications. We then have a chamber decomposition of $\mathrm{Mov}_1(X)$ induced by the decomposition for divisors via the bijection of Theorem 1.7. A good way to analyze the chambers is to compare the behavior of the two functions $\mathfrak{M}$ and $\widehat{\mathrm{vol}}$ restricted to $\mathrm{Mov}_1(X)$.

- By Theorem 1.7, a curve class in the interior of $\mathrm{Mov}_1(X)$ is the $(n-1)$-positive product of a big divisor class L and $\mathfrak{M}(\alpha) = \mathrm{vol}(L)$. Using the birational invariance of the volume for divisors, we see that $\mathfrak{M}$ is also invariant under small $\mathbb{Q}$-factorial modifications.
- Using the Zariski decomposition of [Lehmann and Xiao 2016], the movable cone of curves admits a "chamber structure" as a union of the complete intersection cones from small $\mathbb{Q}$-factorial modifications. However, $\widehat{\mathrm{vol}}$ is not invariant under small $\mathbb{Q}$-factorial modifications but changes to reflect the differing structure of the pseudoeffective cone of curves.

Theorem 1.11 shows that $\widehat{\mathrm{vol}}$ reaches its minimum value $\widehat{\mathrm{vol}}(\alpha) = \mathfrak{M}(\alpha)$ precisely on the complete intersection cone of X, and then increases on the chambers corresponding to birational models of X. In this way $\widehat{\mathrm{vol}}$ and $\mathfrak{M}$ are the right tools for understanding the birational geometry of curves on Mori dream spaces.

Toric varieties. Suppose that X is a simplicial projective toric variety of dimension n defined by a fan Σ. A class α in the interior of the movable cone of curves corresponds to a positive Minkowski weight on the rays of Σ. A fundamental theorem of Minkowski attaches to such a weight a polytope P_α whose facet normals are the rays of Σ and whose facet volumes are determined by the weights. In fact, Minkowski's construction exactly corresponds to the bijection of Theorem 1.7.

Lemma 1.13. *If L denotes the big and movable divisor class corresponding to the polytope P_α then $\langle L^{n-1}\rangle = \alpha$. Thus $\mathfrak{M}(\alpha) = n!\,\mathrm{vol}(P_\alpha)$.*

When α happens to be in the complete intersection cone, this quantity also agrees with $\widehat{\mathrm{vol}}(\alpha)$. In the toric setting, properties of $\mathfrak{M}$ can be interpreted via the classical theory of convex bodies, using constructions such as Blaschke addition and the Kneser–Süss inequality (see [Lehmann and Xiao 2017] for more details).

Further applications. The refined structure of the movable cone is not only important to study positivity functions for curves, but it should also have other applications. We briefly mention two areas for further study (which will not be addressed in the body of the paper).

The first is the study of moduli of vector bundles. Recently, the papers [Greb et al. 2016b; 2016c; 2016d; 2019] discussed some obstructions to generalizing the theory of slope-stability from surfaces to varieties of arbitrary dimension. Traditionally one uses stability conditions defined by H^{n-1} for an ample divisor H, but the walls are no longer linear in H. As discussed in [Greb et al. 2016d] the situation is improved by working in $\mathrm{CI}_1(X)$. Since this cone is not convex, it seems that a thorough understanding of Theorem 1.7 and of stability conditions constructed via movable curve classes (as in [Greb et al. 2016a]) will be helpful for filling out this picture. There are also some situations where one obtains a nice chamber structure of $\mathrm{Mov}_1(X)$ using stability conditions (see for example [Neumann 2010]), and it would be interesting to see the geometric input provided by the corresponding decomposition of $\mathrm{Mov}^1(X)$.

Another area is the geometry of curves on rationally connected varieties. The original proof of boundedness of smooth Fano varieties by [Campana 1992; Kollár et al. 1992] relied on constructing chains of rational curves and controlling the degree against an ample divisor. Such constructions also have interesting interaction with the volume function of curves (see for example Proposition 6.2). By considering the volume of connecting rational chains, one obtains a "birational" variant of boundedness problems which is interesting for arbitrary rationally connected varieties. See [Lehmann and Xiao 2016] for a more in-depth discussion.

Outline of the paper. In this paper we will work with projective varieties over $\mathbb{C}$, but related results can be also adjusted to arbitrary algebraically closed fields and compact Kähler manifolds. We give a general framework for this extension in Section 2.

In Section 2 we briefly recall the general convexity and duality framework in [Lehmann and Xiao 2016], and explain how the proofs can be adjusted to arbitrary algebraically closed fields and compact Kähler manifolds. In Section 3, we give a refined structure of the movable cone of curves and generalize several results on big and nef divisors to big and movable divisors. Section 4 discusses toric varieties, showing some relationships with convex geometry. Section 5 compares the complete intersection and movable cone of curves. In Section 6 we compare the (weighted) mobility functions and $\widehat{\mathrm{vol}}$, $\mathfrak{M}$, finishing the proof of the main results.

2. Preliminaries

Positivity. In this section, we first fix some notations over a projective variety X:

- $N^1(X)$: the real vector space of numerical classes of divisors.
- $N_1(X)$: the real vector space of numerical classes of curves.
- $\overline{\mathrm{Eff}}^1(X)$: the cone of pseudoeffective divisor classes.
- $\mathrm{Nef}^1(X)$: the cone of nef divisor classes.
- $\mathrm{Mov}^1(X)$: the cone of movable divisor classes.
- $\overline{\mathrm{Eff}}_1(X)$: the cone of pseudoeffective curve classes.
- $\mathrm{Mov}_1(X)$: the cone of movable curve classes, equivalently by [Boucksom et al. 2013] the dual of $\overline{\mathrm{Eff}}^1(X)$.
- $\mathrm{CI}_1(X)$: the closure of the set of all curve classes of the form A^{n-1} for an ample divisor A.

With only a few exceptions, capital letters A, B, D, L will denote $\mathbb{R}$-Cartier divisor classes and Greek letters α, β, γ will denote curve classes. For two curve classes α, β, we write $\alpha \succeq \beta$ and $\alpha \preceq \beta$ to denote that $\alpha - \beta$ and $\beta - \alpha$, respectively, belong to $\overline{\mathrm{Eff}}_1(X)$. We will do similarly for divisor classes, or two elements of a cone $\mathfrak{C}$ if the cone is understood.

We will use the notation $\langle - \rangle$ for the positive product as in [Boucksom 2002a; Boucksom et al. 2009; 2013]. Let us recall briefly recall its definition. Let X be a projective manifold (or compact Kähler manifold) of dimension n, and let $L_1, \dots, L_r$ be big $(1, 1)$ classes. Then

$$\langle L_1 \cdots L_r \rangle := \lim_{m\to\infty} \mu_{m*}(\hat{A}_1 \cdots \hat{A}_r),$$

where $\mu_m : X_m \to X$ is a suitable sequence of Fujita approximations such that the limit class has the most positivity (see [Boucksom et al. 2009; 2013] for more details). Note that μ_m satisfies $\mu_m^* L_i = \hat{A}_{i,m} + E_{i,m}$ for some effective divisor class $E_{i,m}$ and big nef class $\hat{A}_{i,m}$ such that $\hat{A}_{i,m}^n \to \mathrm{vol}(L_i)$. We make a few remarks on this construction for singular projective varieties. Suppose that X has dimension n. Then $N_{n-1}(X)$ denotes the vector space of $\mathbb{R}$-classes of Weil divisors up to numerical equivalence as in [Fulton 1984, Chapter 19]. In this setting, the first and $(n-1)$-st positive product should be interpreted respectively as maps $\overline{\mathrm{Eff}}^1(X) \to N_{n-1}(X)$ and $\overline{\mathrm{Eff}}^1(X)^{\times n-1} \to \mathrm{Mov}_1(X)$. We will also let $P_\sigma(L)$ denote the positive part in this sense — that is, pull back L to better and better Fujita approximations, take its positive part, and push the numerical class forward to X as a numerical Weil divisor class. With these conventions, we still have the crucial result of [Boucksom et al. 2009; Lazarsfeld and Mustaţă 2009] that the derivative of the volume is controlled by intersecting against the positive part.

We define the movable cone of divisors $\mathrm{Mov}^1(X)$ to be the subset of $\overline{\mathrm{Eff}}^1(X)$ consisting of divisor classes L such that $N_\sigma(L) = 0$ and $P_\sigma(L) = L \cap [X] \in N_{n-1}(X)$. On any projective variety, by [Fulton 1984, Example 19.3.3] capping with X defines an injective linear map $N^1(X) \to N_{n-1}(X)$. Thus if $D, L \in \mathrm{Mov}^1(X)$ have the same positive part in $N_{n-1}(X)$, then by the injectivity of the capping map we must have $D = L$.

To extend our results (especially the results in Section 3) to arbitrary compact Kähler manifolds, we need to deal with transcendental objects which are not given by divisors or curves. Let X be a compact Kähler manifold of dimension n. By analogue with the projective situation, we need to deal with the following spaces and positive cones:

- $H^{1,1}_{\mathrm{BC}}(X, \mathbb{R})$: the real Bott–Chern cohomology group of bidegree $(1, 1)$.
- $H^{n-1,n-1}_{\mathrm{BC}}(X, \mathbb{R})$: the real Bott–Chern cohomology group of bidegree $(n-1, n-1)$.
- $\mathcal{N}(X)$: the cone of pseudoeffective $(n-1, n-1)$-classes.
- $\mathcal{M}(X)$: the cone of movable $(n-1, n-1)$-classes.
- $\overline{\mathcal{K}}(X)$: the cone of nef $(1, 1)$-classes, equivalently the closure of the Kähler cone.
- $\mathcal{E}(X)$: the cone of pseudoeffective $(1, 1)$-classes.

Recall that we call a Bott–Chern class pseudoeffective if it contains a d-closed positive current, and call an $(n-1, n-1)$-class movable if it is contained in the closure of the cone generated by the classes of the form $\mu_*(\tilde{\omega}_1 \wedge \cdots \wedge \tilde{\omega}_{n-1})$ where $\mu : \tilde{X} \to X$ is a modification and $\tilde{\omega}_1, \ldots, \tilde{\omega}_{n-1}$ are Kähler metrics on $\tilde{X}$. For the basic theory of positive currents, we refer the reader to [Demailly 2012].

Fields of characteristic p. Almost all the results in the paper will hold for smooth varieties over an arbitrary algebraically closed field. The necessary technical generalizations are verified in the following references:

- The existence of Fujita approximations over an arbitrary algebraically closed field is proved in [Takagi 2007].
- The basic properties of the σ-decomposition in positive characteristic are considered in [Mustaţă 2013].
- The results of [Cutkosky 2015] lay the foundations of the theory of positive products and volumes over an arbitrary field.
- [Fulger and Lehmann 2017b] describes how to extend [Boucksom et al. 2013] and most of the results of [Boucksom et al. 2009] over an arbitrary algebraically closed field. In particular the description of the derivative of the volume function in [Boucksom et al. 2009, Theorem A] holds for smooth varieties in any characteristic.

Compact Kähler manifolds. The following results enable us to extend our results in Section 3 and Section 5 to arbitrary compact hyperkähler manifolds and projective manifolds.

- The theory of positive intersection products for pseudoeffective (1, 1)-classes has been developed by [Boucksom 2002a; Boucksom et al. 2010; 2013].
- Divisorial Zariski decomposition for pseudoeffective (1, 1)-classes has been studied in [Boucksom 2004; Boucksom et al. 2013].
- By [Boucksom et al. 2013, Theorem 10.12] and [Nyström and Boucksom 2016], the transcendental analogues of the results in [Boucksom et al. 2009; 2013] are true for compact hyperkähler manifolds and projective manifolds. In particular, we have the cone duality $\mathcal{E}^* = \mathcal{M}$ and the description of the derivative of the volume for pseudoeffective (1, 1)-classes.

Polar transforms. As explained in the introduction, our results use convex analysis, and in particular a Legendre–Fenchel type transform for functions defined on a cone. We briefly recall some definitions and results from [Lehmann and Xiao 2016] which will be used to study the function $\mathfrak{M}$.

Duality transforms. Let V be a finite-dimensional $\mathbb{R}$-vector space of dimension n, and let V^* be its dual. We denote the pairing of $w^* \in V^*$ and $v \in V$ by $w^* \cdot v$. Let $\mathfrak{C} \subset V$ be a proper closed convex cone of full dimension and let $\mathfrak{C}^* \subset V^*$ denote the dual cone of $\mathfrak{C}$. We let $\mathrm{HConc}_s(\mathfrak{C})$ denote the collection of functions $f : \mathfrak{C} \to \mathbb{R}$ satisfying:

- f is upper-semicontinuous and homogeneous of weight $s > 1$.
- f is strictly positive in the interior of $\mathfrak{C}$ (and hence nonnegative on $\mathfrak{C}$).
- f is s-concave: for any $v, x \in \mathfrak{C}$ we have $f(v)^{1/s} + f(x)^{1/s} \le f(v+x)^{1/s}$.

The polar transform $\mathcal{H}$ associates to a function $f \in \mathrm{HConc}_s(\mathfrak{C})$ the function $\mathcal{H}f : \mathfrak{C}^* \to \mathbb{R}$ defined as

$$\mathcal{H}f(w^*) := \inf_{v \in \mathfrak{C}^\circ} \left(\frac{w^* \cdot v}{f(v)^{1/s}} \right)^{s/(s-1)}.$$

The definition is unchanged if we instead vary v over all elements of $\mathfrak{C}$ where f is positive. It is not hard to see that $\mathcal{H}^2 f = f$ for any $f \in \mathrm{HConc}_s(\mathfrak{C})$.

It will be crucial to understand which points obtain the infimum in the definition of $\mathcal{H}f$.

Definition 2.1. Let $f \in \mathrm{HConc}_s(\mathfrak{C})$. For any $w^* \in \mathfrak{C}^*$, we define G_{w^*} to be the set of all $v \in \mathfrak{C}$ which satisfy $f(v) > 0$ and which achieve the infimum in the definition of $\mathcal{H}f(w^*)$, so that

$$\mathcal{H}f(w^*) = \left(\frac{w^* \cdot v}{f(v)^{1/s}} \right)^{s/(s-1)}.$$

Remark 2.2. The set G_{w^*} is the analogue of super-gradients of concave functions. In particular, we know the differential of $\mathcal{H}f$ at w^* lies in G_{w^*} if $\mathcal{H}f$ is differentiable.

We next identify the collection of points where f is controlled by $\mathcal{H}$.

Definition 2.3. Let $f \in \mathrm{HConc}_s(\mathfrak{C})$. We define $\mathfrak{C}_f$ to be the set of all $v \in \mathfrak{C}$ such that $v \in G_{w^*}$ for some $w^* \in \mathfrak{C}$ satisfying $\mathcal{H}f(w^*) > 0$.

We say that $f \in \mathrm{HConc}_s(\mathfrak{C})$ is differentiable if it is $\mathcal{C}^1$ on $\mathfrak{C}^\circ$. In this case we define the function

$$D : \mathfrak{C}^\circ \to V^* \quad \text{by} \quad v \mapsto \frac{df(v)}{s}.$$

We will need to understand the behavior of the derivative along the boundary.

Definition 2.4. We say that $f \in \mathrm{HConc}_s(\mathfrak{C})$ is $+$-differentiable if f is $\mathcal{C}^1$ on $\mathfrak{C}^\circ$ and the derivative on $\mathfrak{C}^\circ$ extends to a continuous function on all of $\mathfrak{C}_f$.

Remark 2.5. For $+$-differentiable functions f, we define the function $D : \mathfrak{C}_f \to V^*$ by extending continuously from $\mathfrak{C}^\circ$.

Teissier proportionality and strict log concavity. In [Lehmann and Xiao 2016], we gave some conditions which are equivalent to the strict log concavity.

Definition 2.6. Let $f \in \mathrm{HConc}_s(\mathfrak{C})$ be $+$-differentiable and let $\mathfrak{C}_T$ be a nonempty subcone of $\mathfrak{C}_f$. We say that f satisfies Teissier proportionality with respect to $\mathfrak{C}_T$ if for any $v, x \in \mathfrak{C}_T$ satisfying

$$D(v) \cdot x = f(v)^{s-1/s} f(x)^{1/s}$$

we have that v and x are proportional.

Note that we do not assume that $\mathfrak{C}_T$ is convex — indeed, in examples it is important to avoid this condition. However, since f is defined on the convex hull of $\mathfrak{C}_T$, we can (somewhat abusively) discuss the strict log concavity of $f|_{\mathfrak{C}_T}$:

Definition 2.7. Let $\mathfrak{C}' \subset \mathfrak{C}$ be a (possibly nonconvex) subcone. We say that f is strictly log concave on $\mathfrak{C}'$ if

$$f(v)^{1/s} + f(x)^{1/s} < f(v+x)^{1/s}$$

holds whenever $v, x \in \mathfrak{C}'$ are not proportional. Note that this definition makes sense even when $\mathfrak{C}'$ is not itself convex.

Theorem 2.8 [Lehmann and Xiao 2016, Theorem 4.12]. *Let $f \in \mathrm{HConc}_s(\mathfrak{C})$ be $+$-differentiable. For any nonempty subcone $\mathfrak{C}_T$ of $\mathfrak{C}_f$, consider the following conditions*:

(1) *The restriction $f|_{\mathfrak{C}_T}$ is strictly log concave* (*in the sense defined above*).

(2) *f satisfies Teissier proportionality with respect to $\mathfrak{C}_T$.*

(3) *The restriction of D to $\mathfrak{C}_T$ is injective.*

Then we have (1) $\Longrightarrow$ (2) $\Longrightarrow$ (3). *If $\mathfrak{C}_T$ is convex, then we have* (2) $\Longrightarrow$ (1). *If $\mathfrak{C}_T$ is an open subcone, then we have* (3) $\Longrightarrow$ (1).

Sublinear boundary conditions. Under certain conditions we can control the behavior of $\mathcal{H}f$ near the boundary, and thus obtain the continuity.

Definition 2.9. Let $f \in \mathrm{HConc}_s(\mathfrak{C})$ and let $\alpha \in (0, 1)$. We say that f satisfies the sublinear boundary condition of order α if for any nonzero v on the boundary of $\mathfrak{C}$ and for any x in the interior of $\mathfrak{C}$, there exists a constant $C := C(v, x) > 0$ such that $f(v+\epsilon x)^{1/s} \geq C\epsilon^{\alpha}$.

Note that the condition is always satisfied at v if $f(v) > 0$. Furthermore, the condition is satisfied for any v, x with $\alpha = 1$ by homogeneity and log-concavity, so the crucial question is whether we can decrease α slightly.

Using this sublinear condition, we get the vanishing of $\mathcal{H}f$ along the boundary.

Proposition 2.10 [Lehmann and Xiao 2016, Proposition 4.21]. *Let $f \in \mathrm{HConc}_s(\mathfrak{C})$ satisfy the sublinear boundary condition of order α. Then $\mathcal{H}f$ vanishes along the boundary. As a consequence, $\mathcal{H}f$ extends to a continuous function over V^* by setting $\mathcal{H}f = 0$ outside $\mathfrak{C}^*$.*

Remark 2.11. If f satisfies the sublinear condition, then $\mathfrak{C}^*_{\mathcal{H}f} = \mathfrak{C}^{*\circ}$.

Formal Zariski decompositions. The Legendre–Fenchel transform relates the strict concavity of a function to the differentiability of its transform. The transform $\mathcal{H}$ will play the same role in our situation; however, one needs to interpret the strict concavity slightly differently. We will encapsulate this property using the notion of a Zariski decomposition.

Definition 2.12. Let $f \in \mathrm{HConc}_s(\mathfrak{C})$ and let $U \subset \mathfrak{C}$ be a nonempty subcone. We say that f admits a strong Zariski decomposition with respect to U if:

(1) For every $v \in \mathfrak{C}_f$ there are unique elements $p_v \in U$ and $n_v \in \mathfrak{C}$ satisfying

$$v = p_v + n_v \quad \text{and} \quad f(v) = f(p_v).$$

We call the expression $v = p_v + n_v$ the Zariski decomposition of v, and call p_v the positive part and n_v the negative part of v.

(2) For any $v, w \in \mathfrak{C}_f$ satisfying $v + w \in \mathfrak{C}_f$ we have

$$f(v)^{1/s} + f(w)^{1/s} \leq f(v+w)^{1/s}$$

with equality only if p_v and p_w are proportional.

In [Lehmann and Xiao 2016, Theorem 4.3], we proved the following theorem linking the existence of Zariski decomposition structure with differentiability.

Theorem 2.13. *Let $f \in \mathrm{HConc}_s(\mathfrak{C})$. Then we have the following results*:

- *If f is $+$-differentiable, then $\mathcal{H}f$ admits a strong Zariski decomposition with respect to the cone $D(\mathfrak{C}_f) \cup \{0\}$.*
- *If $\mathcal{H}f$ admits a strong Zariski decomposition with respect to a cone U, then f is differentiable.*

In the first situation, one can construct the positive part of w^* by choosing any $v \in G_{w^*}$ with $f(v) > 0$ and choosing p_{w^*} to be the unique element of the ray spanned by $D(v)$ with $\mathcal{H}f(p_{w^*}) = \mathcal{H}f(w^*)$.

Under some additional conditions, we can get the continuity of formal Zariski decompositions (see [Lehmann and Xiao 2016, Theorem 4.6]). Note that for the divisorial Zariski decomposition the continuity is already well known due to the concavity of taking positive parts (see e.g., [Boucksom et al. 2009; Küronya and Maclean 2013; Nakayama 2004]).

Theorem 2.14. *Let $f \in \mathrm{HConc}_s(\mathfrak{C})$ be $+$-differentiable. Then the function taking an element $w^* \in \mathfrak{C}^{*\circ}$ to its positive part p_{w^*} is continuous.*

If furthermore $G_v \cup \{0\}$ is a unique ray for every $v \in \mathfrak{C}_f$ and $\mathcal{H}f$ is continuous on all of $\mathfrak{C}^*_{\mathcal{H}f}$, then the Zariski decomposition is continuous on all of $\mathfrak{C}^*_{\mathcal{H}f}$.

Zariski decomposition for curves. In [Lehmann and Xiao 2016], as an application of the above formal Zariski decomposition to the situation

$$\mathfrak{C} = \mathrm{Nef}^1(X), \quad f = \mathrm{vol}, \quad \mathfrak{C}^* = \overline{\mathrm{Eff}}_1(X), \quad \mathcal{H}f = \widehat{\mathrm{vol}},$$

we obtain the Zariski decomposition for curves. The following result is important in the proof of Theorem 1.3.

Definition 2.15. Let X be a projective variety of dimension n and let $\alpha \in \overline{\mathrm{Eff}}_1(X)^\circ$ be a big curve class. Then a Zariski decomposition for α is a decomposition

$$\alpha = B^{n-1} + \gamma$$

where B is a big and nef $\mathbb{R}$-Cartier divisor class, γ is pseudoeffective, and $B \cdot \gamma = 0$. We call B^{n-1} the "positive part" and γ the "negative part" of the decomposition.

Theorem 2.16. *Let X be a projective variety of dimension n and let $\alpha \in \overline{\mathrm{Eff}}_1(X)^\circ$ be a big curve class. Then α admits a unique Zariski decomposition $\alpha = B_\alpha^{n-1} + \gamma$. Furthermore,*

$$\widehat{\mathrm{vol}}(\alpha) = \widehat{\mathrm{vol}}(B_\alpha^{n-1}) = \mathrm{vol}(B_\alpha)$$

and B_α is the unique big and nef divisor class with this property satisfying $B_\alpha^{n-1} \preceq \alpha$. The class B_α depends continuously on α.

Remark 2.17. As explained in [Lehmann and Xiao 2016, Remark 5.1], the above result holds in the Kähler setting — we have a similar decomposition for any interior point of the pseudoeffective $(n-1, n-1)$-cone $\mathcal{N}$.

3. Refined structure of the movable cone

In this section, we study the movable cone of curves and its relationship to the positive product of divisors. A key tool in this study is the following function of [Xiao 2017, Definition 2.2]:

Definition 3.1. Let X be a projective variety of dimension n. For any curve class $\alpha \in \mathrm{Mov}_1(X)$ define

$$\mathfrak{M}(\alpha) = \inf_{L \text{ big divisor class}} \left(\frac{L \cdot \alpha}{\mathrm{vol}(L)^{1/n}} \right)^{n/(n-1)}.$$

We say that a big class L computes $\mathfrak{M}(\alpha)$ if this infimum is achieved by L. When α is a curve class that is not movable, we set $\mathfrak{M}(\alpha) = 0$.

In other words, $\mathfrak{M}$ is the function on $\mathrm{Mov}_1(X)$ defined as the polar transform of the volume function on $\overline{\mathrm{Eff}}^1(X)$. Dually, we can think of the volume function on divisors as the polar transform of $\mathfrak{M}$; this viewpoint allows us to apply the general theory of convexity developed in [Lehmann and Xiao 2016] to vol.

In this section we first prove some new results concerning the volume function for divisors. We will then return to the study of $\mathfrak{M}$ below, where we show that it measures the volume of the "$(n-1)$-st root" of α.

The volume function on big and movable divisors. We first extend several well-known results on big and nef divisors to big and movable divisors. The key will be an extension of Teissier proportionality theorem for big and nef divisors (see [Lehmann and Xiao 2016; Boucksom et al. 2009]) to big and movable divisors.

Lemma 3.2. *Let X be a projective variety of dimension n. Let L_1 and L_2 be big movable divisor classes. Set s to be the largest real number such that $L_1 - sL_2$ is pseudoeffective. Then*

$$s^n \leq \frac{\mathrm{vol}(L_1)}{\mathrm{vol}(L_2)}$$

with equality if and only if L_1 and L_2 are proportional.

Proof. We first prove the case when X is smooth. Certainly we have $\mathrm{vol}(L_1) \geq \mathrm{vol}(sL_2) = s^n \,\mathrm{vol}(L_2)$. If they are equal, then since sL_2 is movable and $L_1 - sL_2$ is pseudoeffective we get a Zariski decomposition of

$$L_1 = sL_2 + (L_1 - sL_2)$$

in the sense of [Fulger and Lehmann 2017b]. By [Fulger and Lehmann 2017b, Proposition 5.3], this decomposition coincides with the numerical version of the σ-decomposition of [Nakayama 2004] so that $P_\sigma(L_1) = sL_2$. Since L_1 is movable, we obtain equality $L_1 = sL_2$.

For arbitrary X, let $\phi : X' \to X$ be a resolution. The inequality follows by pulling back L_1 and L_2 and replacing them by their positive parts. Indeed using the numerical analogue of [Nakayama 2004, III.1.14 Proposition] we see that $\phi^*L_1 - sP_\sigma(\phi^*L_2)$ is pseudoeffective if and only if $P_\sigma(\phi^*L_1) - sP_\sigma(\phi^*L_2)$ is pseudoeffective, so that s can only go up under this operation. To characterize the equality, recall that if L_1 and L_2 are movable and $P_\sigma(\phi^*L_1) = sP_\sigma(\phi^*L_2)$ as elements of $N_{n-1}(X)$, then $L_1 = sL_2$ as elements of $N^1(X)$ by the injectivity of the capping map. □

Next we prove the Diskant inequality for big and movable divisor classes, generalizing the version for big and nef divisors in [Boucksom et al. 2009].

Proposition 3.3. *Let X be a smooth projective variety of dimension n. Let L_1, L_2 be big and movable divisor classes. Set s_L to be the largest real number such that $L_1 - s_L L_2$ is pseudoeffective. Then*

$$(\langle L_1^{n-1}\rangle \cdot L_2)^{n/(n-1)} - \mathrm{vol}(L_1)\,\mathrm{vol}(L_2)^{1/n-1} \geq ((\langle L_1^{n-1}\rangle \cdot L_2)^{1/n-1} - s_L\,\mathrm{vol}(L_2)^{1/n-1})^n. \qquad (\dagger)$$

Proof. Fix an ample divisor H on X.

For any $\epsilon > 0$, by taking sufficiently good Fujita approximations we may find a birational map $\phi_\epsilon : Y_\epsilon \to X$ and ample divisor classes $A_{1,\epsilon}$ and $A_{2,\epsilon}$ such that

- $\phi_\epsilon^* L_i - A_{i,\epsilon}$ is pseudoeffective for $i = 1, 2$;
- $\mathrm{vol}(A_{i,\epsilon}) > \mathrm{vol}(L_i) - \epsilon$ for $i = 1, 2$;
- $\phi_{\epsilon *} A_{i,\epsilon}$ is in an ϵ-ball around L_i for $i = 1, 2$.

Furthermore:

- By applying the argument of [Fulger and Lehmann 2017b, Theorem 6.22], we may ensure
$$\phi_\epsilon^*(\langle L_1^{n-1}\rangle - \epsilon H^{n-1}) \preceq A_{1,\epsilon}^{n-1} \preceq \phi_\epsilon^*(\langle L_1^{n-1}\rangle + \epsilon H^{n-1}).$$
- Set s_ϵ to be the largest real number such that $A_{1,\epsilon} - s_\epsilon A_{2,\epsilon}$ is pseudoeffective. Then we may ensure that $s_\epsilon < s_L + \epsilon$.

By the Khovanskii–Teissier inequality for nef divisor classes, we have

$$(A_{1,\epsilon}^{n-1} \cdot A_{2,\epsilon})^{n/(n-1)} \geq \mathrm{vol}(A_{1,\epsilon})\,\mathrm{vol}(A_{2,\epsilon})^{1/n-1}.$$

Note that $\langle L^{n-1}\rangle \cdot L_2$ is approximated by $A_{1,\epsilon}^{n-1} \cdot A_{2,\epsilon}$ by the projection formula. Taking a limit as ϵ goes to 0, we see that

$$\langle L_1^{n-1}\rangle \cdot L_2 \geq \mathrm{vol}(L_1)^{n-1/n}\,\mathrm{vol}(L_2)^{1/n}. \qquad (\star)$$

On the other hand, the Diskant inequality for big and nef divisors in [Boucksom et al. 2009, Theorem F] implies that

$$\begin{aligned}(A_{1,\epsilon}^{n-1} \cdot A_{2,\epsilon})^{n/(n-1)} - \mathrm{vol}(A_{1,\epsilon})\,\mathrm{vol}(A_{2,\epsilon})^{1/n-1} &\geq ((A_{1,\epsilon}^{n-1} \cdot A_{2,\epsilon})^{1/n-1} - s_\epsilon\,\mathrm{vol}(A_{2,\epsilon})^{1/n-1})^n \\ &\geq ((A_{1,\epsilon}^{n-1} \cdot A_{2,\epsilon})^{1/n-1} - (s_L + \epsilon)\,\mathrm{vol}(A_{2,\epsilon})^{1/n-1})^n.\end{aligned}$$

Taking a limit as ϵ goes to 0 again, we see that

$$(\langle L_1^{n-1}\rangle \cdot L_2)^{n/(n-1)} - \mathrm{vol}(L_1)\,\mathrm{vol}(L_2)^{1/n-1} \geq ((\langle L_1^{n-1}\rangle \cdot L_2)^{1/n-1} - s_L\,\mathrm{vol}(L_2)^{1/n-1})^n.$$

This finishes the proof of the Diskant inequality for big and movable divisor classes. □

Remark 3.4. As shown in [Lehmann and Xiao 2017, Section 3] and implicitly proved in [Fulger and Lehmann 2017b] (which in turn follows from a result of [Boucksom et al. 2009]), for two big movable divisor classes L_1, L_2, we indeed have $\langle L_1^{n-1}\rangle \cdot L_2 = \langle L_1^{n-1} \cdot L_2\rangle$.

As a corollary of Proposition 3.3, we get:

Proposition 3.5. *Let X be a projective variety of dimension n. Let L_1, L_2 be big and movable divisor classes. Then*

$$\langle L_1^{n-1} \rangle \cdot L_2 \geq \mathrm{vol}(L_1)^{n-1/n} \, \mathrm{vol}(L_2)^{1/n}$$

with equality if and only if L_1 and L_2 are proportional.

Proof. If X is smooth, then the result follows directly from Lemma 3.2, $\star$ and $\dagger$.

Now suppose X is singular. The inequality can be computed by passing to a resolution $\phi : X' \to X$ and replacing L_1 and L_2 by their positive parts, since the left-hand side can only decrease under this operation. To characterize the equality, recall that if L_1 and L_2 are movable and $P_\sigma(\phi^* L_1) = s P_\sigma(\phi^* L_2)$ as elements of $N_{n-1}(X)$, then $L_1 = s L_2$ as elements of $N^1(X)$ by the injectivity of the capping map. □

Remark 3.6. In the analytic setting, applying Proposition 3.5 and the same method as [Lehmann and Xiao 2016], it is not hard to generalize Proposition 3.5 to any number of big and movable divisor classes provided we have sufficient regularity for degenerate Monge–Ampère equations in big classes:

- Let $L_1, \ldots, L_n$ be n big divisor classes over a smooth complex projective variety X, then we have

$$\langle L_1 \cdots L_n \rangle \geq \mathrm{vol}(L_1)^{1/n} \cdots \mathrm{vol}(L_n)^{1/n}$$

  where the equality is obtained if and only if $P_\sigma(L_1), \ldots, P_\sigma(L_n)$ are proportional.

We only need to characterize the equality situation. To see this, we need the fact that the above positive intersection $\langle L_1 \cdots L_n \rangle$ depends only on the positive parts $P_\sigma(L_i)$, which follows from the analytic construction of positive product [Boucksom 2002a, Proposition 3.2.10]. Then by the method in [Lehmann and Xiao 2016] where we apply [Boucksom et al. 2010] or [Demailly et al. 2014, Theorem D], we reduce it to the case of a pair of divisor classes, e.g., we get

$$\langle P_\sigma(L_1)^{n-1} \cdot P_\sigma(L_2) \rangle = \mathrm{vol}(L_1)^{n-1/n} \, \mathrm{vol}(L_2)^{1/n}.$$

By the definition of positive product we always have

$$\langle P_\sigma(L_1)^{n-1} \cdot P_\sigma(L_2) \rangle \geq \langle P_\sigma(L_1)^{n-1} \rangle \cdot P_\sigma(L_2) \geq \mathrm{vol}(L_1)^{n-1/n} \, \mathrm{vol}(L_2)^{1/n},$$

this then implies the equality

$$\langle P_\sigma(L_1)^{n-1} \rangle \cdot P_\sigma(L_2) = \mathrm{vol}(L_1)^{n-1/n} \, \mathrm{vol}(L_2)^{1/n}.$$

By Proposition 3.5, we immediately obtain the desired result. See also [Lehmann and Xiao 2017, Section 7] for an alternative approach.

Corollary 3.7. *Let X be a smooth projective variety of dimension n. Let $\alpha \in \mathrm{Mov}_1(X)$ be a big movable curve class. All big divisor classes L satisfying $\alpha = \langle L^{n-1} \rangle$ have the same positive part $P_\sigma(L)$.*

Proof. Suppose L_1 and L_2 have the same positive product. We have $\mathrm{vol}(L_1) = \langle L_2^{n-1} \rangle \cdot L_1$ so that $\mathrm{vol}(L_1) \geq \mathrm{vol}(L_2)$. By symmetry we obtain the reverse inequality, hence equality everywhere, and we conclude by Proposition 3.5 and the σ-decomposition for smooth varieties. □

As a consequence of Proposition 3.5, we show the strict log concavity of the volume function vol on the cone of big and movable divisors.

Proposition 3.8. *Let X be a projective variety of dimension n. Then the volume function* vol *is strictly log concave on the cone of big and movable divisor classes.*

Proof. Since the big and movable cone is convex and since the derivative of vol is continuous, this follows immediately from Proposition 3.5 and Theorem 2.8. □

As a consequence, we get:

Theorem 3.9. *Let X be a projective variety of dimension n. Then for any two big divisor classes L_1, L_2, the equality*

$$\mathrm{vol}(L_1+L_2)^{1/n} = \mathrm{vol}(L_1)^{1/n} + \mathrm{vol}(L_2)^{1/n}$$

holds if and only if the positive parts $P(L_1)$, $P(L_2)$ are proportional.

It is well known that $\mathrm{vol}(L_1+L_2)^{1/n} \geq \mathrm{vol}(L_1)^{1/n} + \mathrm{vol}(L_2)^{1/n}$, thus the above result give a characterization on the equality case.

Proof. First, we assume the equality holds. Note that $\mathrm{vol}(L_i) = \mathrm{vol}(P(L_i))$ for $i = 1, 2$, then we get

$$\mathrm{vol}(L_1+L_2)^{1/n} \geq \mathrm{vol}(P(L_1)+P(L_2))^{1/n} \geq \mathrm{vol}(P(L_1))^{1/n} + \mathrm{vol}(P(L_2))^{1/n} = \mathrm{vol}(L_1)^{1/n} + \mathrm{vol}(L_2)^{1/n}.$$

The equality assumption implies that $\mathrm{vol}(P(L_1)+P(L_2))^{1/n} = \mathrm{vol}(P(L_1))^{1/n} + \mathrm{vol}(P(L_2))^{1/n}$, then by Proposition 3.8 the positive parts $P(L_1)$, $P(L_2)$ are proportional.

Next we assume that the positive parts $P(L_1)$, $P(L_2)$ are proportional. We claim that $P(L_1+L_2) = P(L_1)+P(L_2)$. With this claim, it is easy to see the equality for volumes holds. Next we prove the claim. By the divisorial Zariski decomposition, we have two decompositions for L_1+L_2:

$$L_1+L_2 = P(L_1+L_2)+N(L_1+L_2) = P(L_1)+P(L_2)+N(L_1)+N(L_2).$$

Since $P(L_1)$, $P(L_2)$ are proportional, the orthogonality estimate in the divisorial Zariski decomposition implies

$$\langle (P(L_1)+P(L_2))^{n-1}\rangle \cdot (N(L_1)+N(L_2)) = 0.$$

Multiplying by $\langle (P(L_1)+P(L_2))^{n-1}\rangle$ in the two decompositions of L_1+L_2, we get

$$\langle (P(L_1)+P(L_2))^{n}\rangle \geq P(L_1+L_2)\cdot \langle (P(L_1)+P(L_2))^{n-1}\rangle.$$

By the Khovanski–Teissier inequality, this yields that $\mathrm{vol}(P(L_1)+P(L_2)) \geq \mathrm{vol}(P(L_1+L_2))$. However, we always have $\mathrm{vol}(P(L_1+L_2)) \geq \mathrm{vol}(P(L_1)+P(L_2))$, thus the equality holds everywhere. In particular, Proposition 3.5 implies that $P(L_1+L_2) = P(L_1)+P(L_2)$, finishing the proof of our claim. □

***The function* $\mathfrak{M}$.** We now return to the study of the function $\mathfrak{M}$. We are in the situation:

$$\mathfrak{C} = \overline{\mathrm{Eff}}^1(X), \quad f = \mathrm{vol}, \quad \mathfrak{C}^* = \mathrm{Mov}_1(X), \quad \mathcal{H}f = \mathfrak{M}.$$

Note that $\mathfrak{C}^* = \mathrm{Mov}_1(X)$ follows from the main result of [Boucksom et al. 2013].

As preparation for using the polar transform theory, we recall the analytic properties of the volume function for divisors on smooth varieties. By [Boucksom et al. 2009] the volume function on the pseudoeffective cone of divisors is differentiable on the big cone (with $D(L) = \langle L^{n-1} \rangle$). In the notation of Definition 2.3 the cone $\overline{\mathrm{Eff}}^1(X)_{\mathrm{vol}}$ coincides with the big cone, so that vol is $+$-differentiable. The volume function is n-concave, and is strictly n-concave on the big and movable cone by Proposition 3.8. Furthermore, it admits a strong Zariski decomposition with respect to the movable cone of divisors using the σ-decomposition of [Nakayama 2004] and Proposition 3.8.

Remark 3.10. Note that if X is not smooth (or at least $\mathbb{Q}$-factorial), then it is unclear whether vol admits a Zariski decomposition structure with respect to the cone of movable divisors. For this reason, we will focus on smooth varieties in this section. See Remark 3.24 for more details.

Our first task is to understand the behavior of $\mathfrak{M}$ on the boundary of the movable cone of curves. Note that vol does not satisfy a sublinear condition, so that $\mathfrak{M}$ may not vanish on the boundary of $\mathrm{Mov}_1(X)$.

Lemma 3.11. *Let X be a smooth projective variety of dimension n and let α be a movable curve class. Then $\mathfrak{M}(\alpha) = 0$ if and only if α has vanishing intersection against a nonzero movable divisor class L.*

Proof. We first show that if there exists some nonzero movable divisor class M such that $\alpha \cdot M = 0$ then $\mathfrak{M}(\alpha) = 0$. Fix an ample divisor class A. Note that $M + \epsilon A$ is big and movable for any $\epsilon > 0$. Thus there exists some modification $\mu_\epsilon : Y_\epsilon \to X$ and an ample divisor class A_ϵ on Y_ϵ such that $M + \frac{\epsilon}{2} A = \mu_{\epsilon *} A_\epsilon$. So we can write

$$M + \epsilon A = \mu_{\epsilon *}\big(A_\epsilon + \tfrac{\epsilon}{2} \mu_\epsilon^* A\big),$$

which implies

$$\begin{aligned} \mathrm{vol}(M + \epsilon A) &= \mathrm{vol}\big(\mu_{\epsilon *}\big(A_\epsilon + \tfrac{\epsilon}{2} \mu_\epsilon^* A\big)\big) \\ &\geq \mathrm{vol}\big(A_\epsilon + \tfrac{\epsilon}{2} \mu_\epsilon^* A\big) \\ &\geq n\big(\tfrac{\epsilon}{2} \mu_\epsilon^* A\big)^{n-1} \cdot A_\epsilon \\ &\geq c\epsilon^{n-1} A^{n-1} \cdot M. \end{aligned}$$

This estimate shows that the intersection number

$$\rho_\epsilon = \alpha \cdot \frac{M + \epsilon A}{\mathrm{vol}(M + \epsilon A)^{1/n}}.$$

tends to zero as ϵ tends to zero, and so $\mathfrak{M}(\alpha) = 0$.

Conversely, suppose that $\mathfrak{M}(\alpha) = 0$. From the definition of $\mathfrak{M}(\alpha)$, we can take a sequence of big divisor classes L_k with $\mathrm{vol}(L_k) = 1$ such that

$$\lim_{k \to \infty} (\alpha \cdot L_k)^{n/(n-1)} = \mathfrak{M}(\alpha).$$

Moreover, let $P_\sigma(L_k)$ be the positive part of L_k. Then we have $\mathrm{vol}(P_\sigma(L_k)) = 1$ and

$$\alpha \cdot P_\sigma(L_k) \leq \alpha \cdot L_k$$

since α is movable. Thus we can assume the sequence of big divisor classes L_k is movable in the beginning.

Fix an ample divisor A of volume 1, and consider the classes $L_k/(A^{n-1} \cdot L_k)$. These lie in a compact slice of the movable cone, so they must have a nonzero movable accumulation point L, which without loss of generality we may assume is a limit.

Choose a modification $\mu_\epsilon : Y_\epsilon \to X$ and an ample divisor class $A_{\epsilon,k}$ on Y such that

$$A_{\epsilon,k} \leq \mu_\epsilon^* L_k, \quad \mathrm{vol}(A_{\epsilon,k}) > \mathrm{vol}(L_k) - \epsilon$$

Then

$$L_k \cdot A^{n-1} \geq A_{\epsilon,k} \cdot \mu_\epsilon^* A^{n-1} \geq \mathrm{vol}(A_{\epsilon,k})^{1/n}$$

by the Khovanskii–Teissier inequality. Taking a limit over all ϵ, we find $L_k \cdot A^{n-1} \geq \mathrm{vol}(L_k)^{1/n}$. Thus

$$L \cdot \alpha = \lim_{k \to \infty} \frac{L_k \cdot \alpha}{L_k \cdot A^{n-1}} \leq \mathfrak{M}(\alpha)^{n-1/n} = 0. \qquad \square$$

Example 3.12. Note that a movable curve class α with positive $\mathfrak{M}$ need not lie in the interior of the movable cone of curves. A simple example is when X is the blow-up of $\mathbb{P}^2$ at one point, H denotes the pullback of the hyperplane class. For surfaces the functions $\mathfrak{M}$ and vol coincide, so $\mathfrak{M}(H) = 1$ even though H is on the boundary of $\mathrm{Mov}_1(X) = \mathrm{Nef}^1(X)$.

It is also possible for a big movable curve class α to have $\mathfrak{M}(\alpha) = 0$. This occurs for the projective bundle $X = \mathbb{P}_{\mathbb{P}^1}(\mathcal{O} \oplus \mathcal{O} \oplus \mathcal{O}(-1))$. There are two natural divisor classes on X: the class f of the fibers of the projective bundle and the class ξ of the sheaf $\mathcal{O}_{X/\mathbb{P}^1}(1)$. Using for example [Fulger 2011, Theorem 1.1] and [Fulger and Lehmann 2017b, Proposition 7.1], one sees that f and ξ generate the algebraic cohomology classes with the relations $f^2 = 0$, $\xi^2 f = -\xi^3 = 1$ and that $\mathrm{Mov}^1(X) = \langle f, \xi \rangle$ and $\mathrm{Mov}_1(X) = \langle \xi f, \xi^2 + \xi f \rangle$. We see that the big and movable curve class $\xi^2 + \xi f$ has vanishing intersection against the movable divisor ξ so that $\mathfrak{M}(\xi^2 + \xi f) = 0$ by Lemma 3.11.

Remark 3.13. Another perspective on Lemma 3.11 is provided by the numerical dimension of [Nakayama 2004; Boucksom 2004]. On a smooth variety the following conditions are equivalent for a class $L \in \mathrm{Mov}^1(X)$. (They both correspond to the nonvanishing of the numerical dimension.)

- Fix an ample divisor class A. For any big class D, there is a positive constant C such that $Ct^{n-1} < \mathrm{vol}(L + tA)$ for all $t > 0$.
- $L \neq 0$.

In particular, this implies that vol satisfies the sublinear boundary condition of order $n - 1/n$ when restricted to the movable cone, and this fact can be used in the previous proof. A variant of this statement in characteristic p is proved by [Cascini et al. 2014].

In many ways it is most natural to define $\mathfrak{M}$ using the movable cone of divisors instead of the pseudoeffective cone of divisors. Conceptually, this coheres with the fact that the polar transform can be calculated using the positive part of a Zariski decomposition. Recall that the positive part $P_\sigma(L)$ of a pseudoeffective divisor L has $P_\sigma(L) \preceq L$ and $\mathrm{vol}(P_\sigma(L)) = \mathrm{vol}(L)$. Arguing as in Lemma 3.11 by taking positive parts, we see that for any $\alpha \in \mathrm{Mov}_1(X)$ we have

$$\mathfrak{M}(\alpha) = \inf_{D \text{ big and movable}} \left(\frac{D \cdot \alpha}{\mathrm{vol}(D)^{1/n}} \right)^{n/(n-1)}.$$

Thus for X smooth it is perhaps better to consider the following polar transform:

$$\mathfrak{C} = \mathrm{Mov}^1(X), \quad f = \mathrm{vol}, \quad \mathfrak{C}^* = \mathrm{Mov}^1(X)^*, \quad \mathcal{H}f = \mathfrak{M}'.$$

Since vol satisfies a sublinear condition on $\mathrm{Mov}^1(X)$, the function $\mathfrak{M}'$ is strictly positive exactly in $\mathrm{Mov}^1(X)^{*\circ}$ and extends to a continuous function over $N_1(X)$. The relationship between the two functions is given by

$$\mathfrak{M}'|_{\mathrm{Mov}_1(X)} = \mathfrak{M};$$

this follows immediately from the description for $\mathfrak{M}$ earlier in this paragraph. In fact by Theorem 2.13 $\mathfrak{M}'$ admits a strong Zariski decomposition. Using the interpretation of positive parts via derivatives as in Theorem 2.13, the results of [Boucksom et al. 2009; Lazarsfeld and Mustaţă 2009] show that the positive parts for the Zariski decomposition of $\mathfrak{M}'$ lie in $\mathrm{Mov}_1(X)$. In this way one can think of $\mathfrak{M}$ as the "Zariski projection" of $\mathfrak{M}'$.

Note one important consequence of this perspective: Lemma 3.11 shows that the subcone of $\mathrm{Mov}_1(X)$ where $\mathfrak{M}$ is positive lies in the interior of $\mathrm{Mov}^1(X)^*$. Thus this region agrees with $\mathrm{Mov}_1(X)_{\mathfrak{M}}$ and $\mathfrak{M}$ extends to a differentiable function on an open set containing this cone by applying Theorem 2.13. In particular $\mathfrak{M}$ is $+$-differentiable and continuous on $\mathrm{Mov}_1(X)$.

We next prove a refined structure of the movable cone of curves. Recall that by [Boucksom et al. 2013] the movable cone of curves $\mathrm{Mov}_1(X)$ is generated by the $(n-1)$-self positive products of big divisors. In other words, any curve class in the interior of $\mathrm{Mov}_1(X)$ is a *convex combination* of such positive products. We show that $\mathrm{Mov}_1(X)$ actually coincides with the closure of such products (which naturally form a cone).

Theorem 3.14. *Let X be a smooth projective variety of dimension n. Then any movable curve class α with $\mathfrak{M}(\alpha) > 0$ has the form*

$$\alpha = \langle L_\alpha^{n-1} \rangle$$

for a unique big and movable divisor class L_α. We then have $\mathfrak{M}(\alpha) = \mathrm{vol}(L_\alpha)$ and any big and movable divisor computing $\mathfrak{M}(\alpha)$ is proportional to L_α.

Proof. Applying Theorem 2.13 to $\mathfrak{M}'$, we get

$$\alpha = D(L_\alpha) + n_\alpha$$

where L_α is a big movable class computing $\mathfrak{M}(\alpha)$ and $n_\alpha \in \mathrm{Mov}^1(X)^*$. As D is the differential of $\mathrm{vol}^{1/n}$ on big and movable divisor classes, we have $D(L_\alpha) = \langle L_\alpha^{n-1} \rangle$. Note that $\mathfrak{M}(\alpha) = \langle L_\alpha^{n-1} \rangle \cdot L_\alpha = \mathrm{vol}(L_\alpha)$.

To finish the proof, we observe that $n_\alpha \in \mathrm{Mov}_1(X)$. This follows since α is movable: by the definition of L_α, for any pseudoeffective divisor class E and $t \geq 0$ we have

$$\frac{\alpha \cdot L_\alpha}{\mathrm{vol}(L_\alpha)^{1/n}} \leq \frac{\alpha \cdot P_\sigma(L_\alpha + tE)}{\mathrm{vol}(L_\alpha + tE)^{1/n}} \leq \frac{\alpha \cdot (L_\alpha + tE)}{\mathrm{vol}(L_\alpha + tE)^{1/n}}$$

with equality at $t = 0$. This then implies

$$n_\alpha \cdot E \geq 0.$$

Thus $n_\alpha \in \mathrm{Mov}_1(X)$. Intersecting against L_α, we have $n_\alpha \cdot L_\alpha = 0$. This shows $n_\alpha = 0$ because L_α is an interior point of $\overline{\mathrm{Eff}}^1(X)$ and $\overline{\mathrm{Eff}}^1(X)^* = \mathrm{Mov}_1(X)$. So we have $\alpha = D(L_\alpha) = \langle L_\alpha^{n-1} \rangle$.

Finally, uniqueness follows from Corollary 3.7. □

We note in passing an immediate consequence:

Corollary 3.15. *Let X be a projective variety of dimension n. Then the rays spanned by classes of irreducible curves which deform to cover X are dense in* $\mathrm{Mov}_1(X)$.

Proof. It suffices to prove this on a resolution of X. By Theorem 3.14 it suffices to show that any class of the form $\langle L^{n-1} \rangle$ for a big divisor L is a limit of rescalings of classes of irreducible curves which deform to cover X. Indeed, we may even assume that L is a $\mathbb{Q}$-Cartier divisor. Then the positive product is a limit of the pushforward of complete intersections of ample divisors on birational models, whence the result. □

We can also describe the boundary of $\mathrm{Mov}_1(X)$, in combination with Lemma 3.11.

Corollary 3.16. *Let X be a smooth projective variety of dimension n. Let α be a movable class with $\mathfrak{M}(\alpha) > 0$ and let L_α be the unique big movable divisor whose positive product is α. Then α is on the boundary of* $\mathrm{Mov}_1(X)$ *if and only if L_α is on the boundary of* $\mathrm{Mov}^1(X)$.

Proof. Note that α is on the boundary of $\mathrm{Mov}_1(X)$ if and only if it has vanishing intersection against a class D lying on an extremal ray of $\overline{\mathrm{Eff}}^1(X)$. Lemma 3.11 shows that in this case D is not movable, so by [Nakayama 2004, Chapter III.1] D is (after rescaling) the class of an integral divisor on X which we continue to call D. By [Boucksom et al. 2009, Proposition 4.8 and Theorem 4.9], the equation $\langle L_\alpha^{n-1} \rangle \cdot D = 0$ holds if and only if $D \in \mathbb{B}_+(L_\alpha)$. Altogether, we see that α is on the boundary of $\mathrm{Mov}_1(X)$ if and only if L_α is on the boundary of $\mathrm{Mov}^1(X)$. □

Arguing using abstract properties of polar transforms just as in [Lehmann and Xiao 2016], the good analytic properties of the volume function for divisors imply most of the other analytic properties of $\mathfrak{M}$.

Theorem 3.17 (see Theorem 2.14, and compare with [Lehmann and Xiao 2016, Theorem 5.6]). *Let X be a smooth projective variety of dimension n. For any movable curve class α with $\mathfrak{M}(\alpha) > 0$, let L_α denote the unique big and movable divisor class satisfying $\langle L_\alpha^{n-1} \rangle = \alpha$. As we vary α in* $\mathrm{Mov}_1(X)_{\mathfrak{M}}$, *$L_\alpha$ depends continuously on α.*

Theorem 3.18 (compare with [Lehmann and Xiao 2016, Theorem 5.11]). *Let X be a smooth projective variety of dimension n. For a curve class $\alpha = \langle L_\alpha^{n-1}\rangle$ in $\mathrm{Mov}_1(X)_{\mathfrak{M}}$ and for an arbitrary curve class $\beta \in N_1(X)$ we have*

$$\frac{d}{dt}\bigg|_{t=0} \mathfrak{M}(\alpha + t\beta) = \frac{n}{n-1} P_\sigma(L_\alpha)\cdot\beta.$$

Theorem 3.19 (see Theorem 2.13, and compare with [Lehmann and Xiao 2016, Theorem 5.10]). *Let X be a smooth projective variety of dimension n. Let α_1, α_2 be two big and movable curve classes in* $\mathrm{Mov}_1(X)_{\mathfrak{M}}$. *Then*

$$\mathfrak{M}(\alpha_1+\alpha_2)^{n-1/n} \geq \mathfrak{M}(\alpha_1)^{n-1/n} + \mathfrak{M}(\alpha_2)^{n-1/n}$$

with equality if and only if α_1 and α_2 are proportional.

Remark 3.20. Theorem 3.19 can be interpreted as an analogue of the Knesser–Süss inequality for polytopes. We clarify this relationship when discussing toric varieties in Section 4.

Another application of the results in this section is the Morse-type bigness criterion for movable curve classes, which is slightly different from [Lehmann and Xiao 2016, Theorem 5.18].

Theorem 3.21. *Let X be a smooth projective variety of dimension n. Let α, β be two curve classes lying in* $\mathrm{Mov}_1(X)_{\mathfrak{M}}$. *Write $\alpha = \langle L_\alpha^{n-1}\rangle$ and $\beta = \langle L_\beta^{n-1}\rangle$ for the unique big and movable divisor classes L_α, L_β given by Theorem 3.14. Then we have*

$$\mathfrak{M}(\alpha-\beta)^{n-1/n} \geq (\mathfrak{M}(\alpha) - nL_\alpha\cdot\beta)\cdot\mathfrak{M}(\alpha)^{-1/n} = (\mathrm{vol}(L_\alpha) - nL_\alpha\cdot\beta)\cdot\mathrm{vol}(L_\alpha)^{-1/n}.$$

In particular, we have

$$\mathfrak{M}(\alpha-\beta) \geq \mathrm{vol}(L_\alpha) - \frac{n^2}{n-1}L_\alpha\cdot\beta$$

and the curve class $\alpha-\beta$ is big whenever $\mathfrak{M}(\alpha) - nL_\alpha\cdot\beta > 0$.

Proof. By [Lehmann and Xiao 2016, Section 4.2] it suffices to prove a Morse-type bigness criterion for the difference of two movable divisor classes. So we need to prove $L - M$ is big whenever

$$\langle L^n\rangle - n\langle L^{n-1}\rangle\cdot M > 0.$$

This is proved (in the Kähler setting) in [Xiao 2018, Theorem 1.1]. □

Remark 3.22. We remark that we cannot extend this Morse-type criterion from big and movable divisors to arbitrary pseudoeffective divisor classes. A very simple construction provides the counter-examples, e.g., the blow up of $\mathbb{P}^2$ (see [Trapani 1995, Example 3.8]).

Combining Theorem 3.14 and Theorem 3.17, we obtain:

Corollary 3.23. *Let X be a smooth projective variety of dimension n. Then*

$$\Phi : \mathrm{Mov}^1(X)_{\mathrm{vol}} \to \mathrm{Mov}_1(X)_{\mathfrak{M}}, \quad L \mapsto \langle L^{n-1}\rangle$$

is a homeomorphism.

Remark 3.24. Modified versions of many of the results in this section hold for singular varieties as well (see Remark 3.10). For example, by similar arguments we can see that any element in the interior of $\mathrm{Mov}_1(X)$ is the positive product of some big divisor class regardless of singularities. Conversely, whenever $\mathfrak{M}$ is $+$-differentiable we obtain a Zariski decomposition structure for vol by Theorem 2.13.

Remark 3.25. All the results above extend to smooth varieties over algebraically closed fields. However, for compact Kähler manifolds some results rely on Demailly's conjecture on the transcendental holomorphic Morse-type inequality, or equivalently, on the extension of the results of [Boucksom et al. 2009] to the Kähler setting. Since the results of [Boucksom et al. 2009] are used in an essential way in the proof of Theorems 3.14 and 3.2 (via the proof of [Fulger and Lehmann 2017b, Proposition 5.3]), the only statement in this section which extends unconditionally to the Kähler setting is Lemma 3.11. However, these conjectures are known if X is a compact hyperkähler manifold or projective manifold (see [Boucksom et al. 2013, Theorem 10.12; Nyström and Boucksom 2016]), so all of our results extend to compact hyperkähler manifolds.

4. Positivity functions on toric varieties

We study the function $\mathfrak{M}$ on toric varieties, showing that it can be interpreted by the underlying special structures. In this section, X will denote a simplicial projective toric variety of dimension n. In terms of notation, X will be defined by a fan Σ in a lattice N with dual lattice M. We let $\{v_i\}$ denote the primitive generators of the rays of Σ and $\{D_i\}$ denote the corresponding classes of T-divisors. Our goal is to interpret the properties of the function $\mathfrak{M}$ in terms of toric geometry.

Positive product on toric varieties. Suppose that L is a big movable divisor class on the toric variety X. Then L naturally defines a (nonlattice) polytope Q_L; if we choose an expression $L = \sum a_i D_i$, then

$$Q_L = \{u \in M_{\mathbb{R}} \mid \langle u, v_i \rangle + a_i \geq 0\}$$

and changing the choice of representative corresponds to a translation of Q_L. Conversely, suppose that Q is a full-dimensional polytope such that the unit normals to the facets of Q form a subset of the rays of Σ. Then Q uniquely determines a big movable divisor class L_Q on X. The divisors in the interior of the movable cone correspond to those polytopes whose facet normals coincide with the rays of Σ.

Given polytopes $Q_1, \ldots, Q_n$, let $V(Q_1, \ldots, Q_n)$ denote the mixed volume of the polytopes. [Boucksom et al. 2009] explains that the positive product of big movable divisors $L_1, \ldots, L_n$ can be interpreted via the mixed volume of the corresponding polytopes:

$$\langle L_1 \cdots L_n \rangle = n! V(Q_1, \ldots, Q_n).$$

The function $\mathfrak{M}$. In this section we use a theorem of Minkowski to describe the function $\mathfrak{M}$. We thank J. Huh for a conversation working out this picture.

Recall that a class $\alpha \in \mathrm{Mov}_1(X)$ defines a nonnegative Minkowski weight on the rays of the fan Σ — that is, an assignment of a positive real number t_i to each vector v_i such that $\sum t_i v_i = 0$. From now on we will identify α with its Minkowski weight. We will need to identify which movable curve classes are positive along a set of rays which span $\mathbb{R}^n$.

Lemma 4.1. *Suppose $\alpha \in \mathrm{Mov}_1(X)$ satisfies $\mathfrak{M}(\alpha) > 0$. Then α is positive along a spanning set of rays of Σ.*

We will soon see that the converse is also true in Theorem 4.2.

Proof. Suppose that there is a hyperplane V which contains every ray of Σ along which α is positive. Since X is projective, Σ has rays on both sides of V. Let D be the effective divisor consisting of the sum over all the primitive generators of rays of Σ not contained in V. It is clear that the polytope defined by D has nonzero projection onto the subspace spanned by $V^{\perp}$, and in particular, that the polytope defined by D is nonzero. Thus the asymptotic growth of sections of mD is at least linear in m, so $P_\sigma(D) \neq 0$ and α has vanishing intersection against a nonzero movable divisor. Lemma 3.11 shows that $\mathfrak{M}(\alpha) = 0$. $\square$

Minkowski's theorem asserts the following. Suppose that $u_1, \dots, u_s$ are unit vectors which span $\mathbb{R}^n$ and that $r_1, \dots, r_s$ are positive real numbers. Then there exists a polytope P with unit normals $u_1, \dots, u_s$ and with corresponding facet volumes $r_1, \dots, r_s$ if and only if the u_i satisfy

$$r_1 u_1 + \cdots + r_s u_s = 0.$$

Moreover, the resulting polytope is unique up to translation. (See [Klain 2004] for a proof which is compatible with the results below.) If a vector u is a unit normal to a facet of P, we will use the notation $\mathrm{vol}(P^u)$ to denote the volume of the facet corresponding to u.

If α is positive on a spanning set of rays, then it canonically defines a polytope (up to translation) via Minkowski's theorem by choosing the vectors u_i to be the unit vectors in the directions v_i and assigning to each the constant

$$r_i = \frac{t_i |v_i|}{(n-1)!}.$$

Note that this is the natural choice of volume for the corresponding facet, as it accounts for

- the discrepancy in length between u_i and v_i, and
- the factor $1/(n-1)!$ relating the volume of an $(n-1)$-simplex to the determinant of its edge vectors.

We denote the corresponding polytope in $M_{\mathbb{R}}$ defined by the theorem of Minkowski by P_α.

Theorem 4.2. *Suppose α is a movable curve class which is positive on a spanning set of rays and let P_α be the corresponding polytope. Then*

$$\mathfrak{M}(\alpha) = n!\,\mathrm{vol}(P_\alpha).$$

Furthermore, the big movable divisor L_α corresponding to the polytope P_α satisfies $\langle L_\alpha^{n-1} \rangle = \alpha$.

Proof. Let $L \in \mathrm{Mov}^1(X)$ be a big movable divisor class and denote the corresponding polytope by Q_L. We claim that the intersection number can be interpreted as a mixed volume:

$$L \cdot \alpha = n! V(P_\alpha^{n-1}, Q_L).$$

To see this, define for a compact convex set K the function $h_K(u) = \sup_{v \in K}\{v \cdot u\}$. Using [Klain 2004, Equation (5)]

$$V(P_\alpha^{n-1}, Q_L) = \frac{1}{n} \sum_{u \text{ a facet of } P_\alpha + Q_L} h_{Q_L}(u) \operatorname{vol}(P_\alpha^u) = \frac{1}{n} \sum_{\text{rays } v_i} \left(\frac{a_i}{|v_i|}\right)\left(\frac{t_i |v_i|}{(n-1)!}\right) = \frac{1}{n!} \sum_{\text{rays } v_i} a_i t_i = \frac{1}{n!} L \cdot \alpha.$$

Note that we actually have equality in the second line because L is big and movable. Recall that by the Brunn–Minkowski inequality

$$V(P_\alpha^{n-1}, Q_L) \geq \operatorname{vol}(P_\alpha)^{n-1/n} \operatorname{vol}(Q_L)^{1/n}$$

with equality only when P_α and Q_L are homothetic. Thus

$$\mathfrak{M}(\alpha) = \inf_{L \text{ big movable class}} \left(\frac{L \cdot \alpha}{\operatorname{vol}(L)^{1/n}}\right)^{n/(n-1)} = \inf_{L \text{ big movable class}} \left(\frac{n! V(P_\alpha^{n-1}, Q_L)}{n!^{1/n} \operatorname{vol}(Q_L)^{1/n}}\right)^{n/(n-1)} \geq n! \operatorname{vol}(P_\alpha).$$

Furthermore, the equality is achieved for divisors L whose polytope is homothetic to P_α, showing the computation of $\mathfrak{M}(\alpha)$. Furthermore, since the divisor L_α defined by the polytope computes $\mathfrak{M}(\alpha)$ we see that $\langle L_\alpha^{n-1} \rangle$ is proportional to α. By computing $\mathfrak{M}$ we deduce the equality:

$$\mathfrak{M}(\langle L_\alpha^{n-1} \rangle) = \operatorname{vol}(L) = n! \operatorname{vol}(P_\alpha) = \mathfrak{M}(\alpha). \qquad \square$$

The previous result shows:

Corollary 4.3. *Let α be a curve class in $\operatorname{Mov}_1(X)_{\mathfrak{M}}$. Then $\alpha \in \operatorname{CI}_1(X)$ if and only if the normal fan to the corresponding polytope P_α is refined by Σ. In this case we have*

$$\widehat{\operatorname{vol}}(\alpha) = n! \operatorname{vol}(P_\alpha).$$

Proof. By the uniqueness in Theorem 3.14, $\alpha \in \operatorname{CI}_1(X)$ if and only if the corresponding divisor L_α as in Theorem 4.2 is big and nef. $\square$

For toric varieties, much of the theory developed in this paper reduces to results from the theory of convex bodies. For example, suppose that we have movable curve classes α_1, α_2. Then the polytope corresponding to $\alpha_1 + \alpha_2$ is (essentially by definition) the Blaschke sum of the polytopes P_{α_1} and P_{α_2}. Thus the inequality

$$\mathfrak{M}(\alpha_1 + \alpha_2)^{n-1/n} \geq \mathfrak{M}(\alpha_1)^{n-1/n} + \mathfrak{M}(\alpha_2)^{n-1/n}$$

of Theorem 3.19 is exactly the Kneser–Süss inequality when interpreted via toric geometry. Similarly, the derivative formula of Theorem 3.18 follows from the theory of mixed volumes. See [Lehmann and Xiao 2017] for more details.

5. Comparing the complete intersection cone and the movable cone

Consider the functions $\widehat{\operatorname{vol}}$ and $\mathfrak{M}$ on the movable cone of curves $\operatorname{Mov}_1(X)$. By their definitions we always have $\widehat{\operatorname{vol}} \geq \mathfrak{M}$ on the movable cone, and [Xiao 2017, Remark 3.1] asks whether one can characterize when equality holds. In this section we show:

Theorem 5.1. *Let X be a smooth projective variety of dimension n and let α be a big and movable class. Then $\widehat{\mathrm{vol}}(\alpha) = \mathfrak{M}(\alpha)$ if and only if $\alpha \in \mathrm{CI}_1(X)$.*

Thus $\widehat{\mathrm{vol}}$ and $\mathfrak{M}$ can be used to distinguish whether a big movable curve class lies in $\mathrm{CI}_1(X)$ or not. This result is important in Section 6.

Proof. If $\alpha = B^{n-1}$ is a complete intersection class, then $\widehat{\mathrm{vol}}(\alpha) = \mathrm{vol}(B) = \mathfrak{M}(\alpha)$. By continuity the equality holds true for any big curve class in $\mathrm{CI}_1(X)$.

Conversely, suppose that α is not in the complete intersection cone. The claim is clearly true if $\mathfrak{M}(\alpha) = 0$, so by Theorem 3.14 it suffices to consider the case when there is a big and movable divisor class L such that $\alpha = \langle L^{n-1} \rangle$. Note that L can not be big and nef since $\alpha \notin \mathrm{CI}_1(X)$.

We prove $\widehat{\mathrm{vol}}(\alpha) > \mathfrak{M}(\alpha)$ by contradiction. First, by the definition of $\widehat{\mathrm{vol}}$ we always have

$$\widehat{\mathrm{vol}}(\langle L^{n-1} \rangle) \geq \mathfrak{M}(\langle L^{n-1} \rangle) = \mathrm{vol}(L).$$

Suppose $\widehat{\mathrm{vol}}(\langle L^{n-1} \rangle) = \mathrm{vol}(L)$. For convenience, we assume $\mathrm{vol}(L) = 1$. By rescaling the positive part of a Zariski decomposition, we find a big and nef divisor class B with $\mathrm{vol}(B) = 1$ such that $\widehat{\mathrm{vol}}(\langle L^{n-1} \rangle) = (\langle L^{n-1} \rangle \cdot B)^{n/(n-1)}$. For the divisor class B we get

$$\langle L^{n-1} \rangle \cdot B = 1 = \mathrm{vol}(L)^{n-1/n} \mathrm{vol}(B)^{1/n}.$$

By Proposition 3.5, this implies L and B are proportional which contradicts the nonnefness of L. Thus we must have $\widehat{\mathrm{vol}}(\langle L^{n-1} \rangle) > \mathrm{vol}(L) = \mathfrak{M}(\langle L^{n-1} \rangle)$. □

We also obtain:

Proposition 5.2. *Let X be a smooth projective variety of dimension n and let α be a big and movable curve class. Then $\alpha \in \mathrm{CI}_1(X)$ if and only if for any birational morphism $\phi : Y \to X$ we have $\widehat{\mathrm{vol}}(\phi^* \alpha) = \widehat{\mathrm{vol}}(\alpha)$.*

Proof. The forward implication is clear. For the reverse implication, we first consider the case when $\mathfrak{M}(\alpha) > 0$. Let L be a big movable divisor class satisfying $\langle L^{n-1} \rangle = \alpha$. Choose a sequence of birational maps $\phi_\epsilon : Y_\epsilon \to X$ and ample divisor classes A_ϵ on Y_ϵ defining an ϵ-Fujita approximation for L. Then $\mathrm{vol}(L) \geq \mathrm{vol}(A_\epsilon) > \mathrm{vol}(L) - \epsilon$ and the classes $\phi_{\epsilon *} A_\epsilon$ limit to L. Note that $A_\epsilon \cdot \phi_\epsilon^* \alpha = \phi_{\epsilon *} A_\epsilon \cdot \alpha$. This implies that for any $\epsilon > 0$ we have

$$\widehat{\mathrm{vol}}(\alpha) = \widehat{\mathrm{vol}}(\phi_\epsilon^* \alpha) \leq \frac{(\alpha \cdot \phi_{\epsilon *} A_\epsilon)^{n/(n-1)}}{\mathrm{vol}(L)^{1/n-1}}.$$

As ϵ shrinks the right-hand side approaches $\mathrm{vol}(L) = \mathfrak{M}(\alpha)$, and we conclude by Theorem 5.1.

Next we consider the case when $\mathfrak{M}(\alpha) = 0$. Choose a class ξ in the interior of $\mathrm{Mov}_1(X)$ and consider the classes $\alpha + \delta\xi$ for $\delta > 0$. The argument above shows that for any $\epsilon > 0$, there is a birational model $\phi_\epsilon : Y_\epsilon \to X$ such that

$$\widehat{\mathrm{vol}}(\phi_\epsilon^*(\alpha + \delta\xi)) < \mathfrak{M}(\alpha + \delta\xi) + \epsilon.$$

But we also have $\widehat{\mathrm{vol}}(\phi_\epsilon^*\alpha) \leq \widehat{\mathrm{vol}}(\phi_\epsilon^*(\alpha+\delta\xi))$ since the pullback of the nef curve class $\delta\xi$ is pseudoeffective. Taking limits as $\epsilon \to 0$, $\delta \to 0$, we see that we can make the volume of the pullback of α arbitrarily small, a contradiction to the assumption and the bigness of α. □

As an illustration of the comparison between $\widehat{\mathrm{vol}}$ and $\mathfrak{M}$, we discuss Mori dream spaces.

Example 5.3. Let X be a Mori dream space. Recall that a small $\mathbb{Q}$-factorial modification (henceforth SQM) $\phi : X \dashrightarrow X'$ is a birational contraction (i.e., does not extract any divisors) defined in codimension 1 such that X' is projective $\mathbb{Q}$-factorial. Hu and Keel [2000] showed that for any SQM the strict transform defines an isomorphism $\phi_* : N^1(X) \to N^1(X')$ which preserves the pseudoeffective and movable cones of divisors. (More generally, any birational contraction induces an injective pullback $\phi^* : N^1(X') \to N^1(X)$ and dually a surjection $\phi_* : N_1(X) \to N_1(X')$.) The SQM structure induces a chamber decomposition of the pseudoeffective and movable cones of divisors.

One would like to see a "dual picture" in $N_1(X)$ of this chamber decomposition. However, it does not seem interesting to simply dualize the divisor decomposition: the resulting cones are no longer pseudoeffective and are described as intersections instead of unions. Motivated by the Zariski decomposition for curves, we define a chamber structure on the movable cone of curves as a union of the complete intersection cones on SQMs.

Note that for each SQM we obtain by duality an isomorphism $\phi_* : N_1(X) \to N_1(X')$ which preserves the movable cone of curves. We claim that the strict transforms of the various complete intersection cones define a chamber structure on $\mathrm{Mov}_1(X)$. More precisely, given any birational contraction $\phi : X \dashrightarrow X'$ with X' normal projective, define

$$\mathrm{CI}_\phi^\circ := \bigcup_{A \text{ ample on } X'} \langle \phi^* A^{n-1} \rangle.$$

Then:

- $\mathrm{Mov}_1(X)$ is the union over all SQMs $\phi : X \dashrightarrow X'$ of $\overline{\mathrm{CI}_\phi^\circ} = \phi_*^{-1}\mathrm{CI}_1(X')$, and the interiors of the $\overline{\mathrm{CI}_\phi^\circ}$ are disjoint.
- The set of classes in $\mathrm{Mov}_1(X)_{\mathfrak{M}}$ is the disjoint union over all birational contractions $\phi : X \dashrightarrow X'$ of the CI_ϕ°.

To see this, first recall that for a pseudoeffective divisor L the σ-decomposition of L and the volume are preserved by ϕ_*. We know that each $\alpha \in \mathrm{Mov}_1(X)_{\mathfrak{M}}$ has the form $\langle L^{n-1} \rangle$ for a unique big and movable divisor L. If $\phi : X \dashrightarrow X'$ denotes the birational canonical model obtained by running the L-MMP, and A denotes the corresponding ample divisor on X', then $\phi_*\alpha = A^{n-1}$ and $\alpha = \langle \phi^* A^{n-1} \rangle$. The various claims now can be deduced from the properties of divisors and the MMP for Mori dream spaces as in [Hu and Keel 2000, 1.11 Proposition].

Since the volume of divisors behaves compatibly with strict transforms of pseudoeffective divisors, the description of ϕ_* above shows that $\mathfrak{M}$ also behaves compatibly with strict transforms of movable curves under an SQM. However, the volume function can change: we may well have $\widehat{\mathrm{vol}}(\phi_*\alpha) \neq \widehat{\mathrm{vol}}(\alpha)$. The reason is that the pseudoeffective cone of curves is also changing as we vary ϕ. In particular, the set

$$C_{\alpha,\phi} := \{\phi_*\alpha - \gamma \mid \gamma \in \overline{\mathrm{Eff}}_1(X')\}$$

will look different as we vary ϕ. Since $\widehat{\mathrm{vol}}$ is the same as the maximum value of $\mathfrak{M}(\beta)$ for $\beta \in C_{\alpha,\phi}$, the volume and Zariski decomposition for a given model will depend on the exact shape of $C_{\alpha,\phi}$.

Remark 5.4. Theorem 5.1 also holds for smooth varieties over any algebraically closed field and for compact hyperkähler manifolds or projective manifolds as explained in Section 2.

6. Comparison between the positivity functions for curves

Asymptotic point counts and $\widehat{\mathrm{vol}}$. In this section we give the proof of the main result, comparing the volume function for pseudoeffective curves with its mobility function. Recall from the introduction what we are trying to show (slightly reordered):

Theorem 6.1. *Let X be a smooth projective variety of dimension n and let $\alpha \in \overline{\mathrm{Eff}}_1(X)$ be a pseudoeffective curve class. Then the following results hold*:

(1) $\widehat{\mathrm{vol}}(\alpha) \leq \mathrm{mob}(\alpha) \leq n!\,\widehat{\mathrm{vol}}(\alpha)$.

(2) *Assume Conjecture 1.4. Then* $\mathrm{mob}(\alpha) = \widehat{\mathrm{vol}}(\alpha)$.

(3) $\widehat{\mathrm{vol}}(\alpha) = \mathrm{wmob}(\alpha)$.

The upper bound in the first part improves the related result [Xiao 2017, Theorem 3.2]. Before giving the proof, we repeat the following estimate of $\widehat{\mathrm{vol}}$ in [Lehmann and Xiao 2016].

Proposition 6.2. *Let X be a smooth projective variety of dimension n. Choose positive integers $\{k_i\}_{i=1}^r$. Suppose that $\alpha \in \mathrm{Mov}_1(X)$ is represented by a family of irreducible curves such that for any collection of general points $x_1, x_2, \ldots, x_r, y$ of X, there is a curve in our family which contains y and contains each x_i with multiplicity $\geq k_i$. Then*

$$\widehat{\mathrm{vol}}(\alpha)^{n-1/n} \geq \mathfrak{M}(\alpha)^{n-1/n} \geq \frac{\sum_i k_i}{r^{1/n}}.$$

This is just a rephrasing of well-known results in birational geometry; see for example [Kollár 1996, V.2.9 Proposition].

Proof. By continuity and rescaling invariance, it suffices to show that if L is a big and movable Cartier divisor class then

$$\left(\sum_{i=1}^r k_i\right)\frac{\mathrm{vol}(L)^{1/n}}{r^{1/n}} \leq L \cdot C.$$

A standard argument (see for example [Lehmann 2016, Example 8.19]) shows that for any $\epsilon > 0$ and any very general points $\{x_i\}_{i=1}^r$ of X there is a positive integer m and a Cartier divisor M numerically equivalent to mL and such that $\mathrm{mult}_{x_i} M \geq mr^{-1/n}\,\mathrm{vol}(L)^{1/n} - \epsilon$ for every i. By the assumption on the family of curves we may find an irreducible curve C with multiplicity $\geq k_i$ at each x_i that is not contained M. Then

$$m(L \cdot C) \geq \sum_{i=1}^{r} k_i \operatorname{mult}_{x_i} M \geq \left(\sum_{i=1}^{r} k_i\right)\left(\frac{m \operatorname{vol}(L)^{1/n}}{r^{1/n}} - \epsilon\right).$$

Divide by m and let ϵ go to 0 to conclude. □

Example 6.3. The most important special case is when α is the class of a family of irreducible curves such that for any two general points of X there is a curve in our family containing them. Proposition 6.2 then shows that $\widehat{\operatorname{vol}}(\alpha) \geq 1$ and $\mathfrak{M}(\alpha) \geq 1$.

We also need to give a formal definition of the mobility count. Its properties are studied in more depth in [Lehmann 2016].

Definition 6.4. Let X be an integral projective variety and let W be a reduced variety. Suppose that $U \subset W \times X$ is a subscheme and let $p : U \to W$ and $s : U \to X$ denote the projection maps. The mobility count $\operatorname{mc}(p)$ of the morphism p is the maximum nonnegative integer b such that the map

$$U \times_W U \times_W \cdots \times_W U \xrightarrow{s \times s \times \cdots \times s} X \times X \times \cdots \times X$$

is dominant, where we have b terms in the product on each side. (If the map is dominant for every positive integer b, we set $\operatorname{mc}(p) = \infty$.)

For $\alpha \in N_1(X)_{\mathbb{Z}}$, the mobility count of α, denoted $\operatorname{mc}(\alpha)$, is defined to be the largest mobility count of any family of effective curves representing α.

The mobility is then defined as

$$\operatorname{mob}(\alpha) = \limsup_{m \to \infty} \frac{\operatorname{mc}(m\alpha)}{m^{n/(n-1)}/n!}.$$

Proof of Theorem 6.1. (1) We compare mob and $\widehat{\operatorname{vol}}$. We first prove the upper bound. By continuity and homogeneity it suffices to prove the upper bound for a class α in the natural sublattice of integral classes $N_1(X)_{\mathbb{Z}}$. Suppose that $p : U \to W$ is a family of curves representing $m\alpha$ of maximal mobility count for a positive integer m. Suppose that a general member of p decomposes into irreducible components $\{C_i\}$; arguing as in [Lehmann 2016, Corollary 4.10], we must have $\operatorname{mc}(p) = \sum_i \operatorname{mc}(U_i)$, where U_i represents the closure of the family of deformations of C_i. We also let β_i denote the numerical class of C_i.

Suppose that $\operatorname{mc}(U_i) > 1$. Then we may apply Proposition 6.2 with all $k_i = 1$ and $r = \operatorname{mc}(U_i) - 1$ to deduce that

$$\widehat{\operatorname{vol}}(\beta_i) \geq \operatorname{mc}(U_i) - 1.$$

If $\operatorname{mc}(U_i) \leq 1$ then Proposition 6.2 does not apply but at least we still know that $\widehat{\operatorname{vol}}(\beta_i) \geq 0 \geq \operatorname{mc}(U_i) - 1$. Fix an ample Cartier divisor A, and note that the number of components C_i is at most $mA \cdot \alpha$. All told, we have

$$\widehat{\operatorname{vol}}(m\alpha) \geq \sum_i \widehat{\operatorname{vol}}(\beta_i) \geq \sum_i (\operatorname{mc}(U_i) - 1) \geq \operatorname{mc}(m\alpha) - mA \cdot \alpha.$$

Thus,

$$\widehat{\mathrm{vol}}(\alpha) = \limsup_{m\to\infty} \frac{\widehat{\mathrm{vol}}(m\alpha)}{m^{n/(n-1)}} \geq \limsup_{m\to\infty} \frac{\mathrm{mc}(m\alpha) - mA\cdot\alpha}{m^{n/(n-1)}} = \frac{\mathrm{mob}(\alpha)}{n!}.$$

The lower bound relies on the Zariski decomposition of curves in Theorem 2.16. By [Lehmann 2016, Example 6.2] we have

$$B^n \leq \mathrm{mob}(B^{n-1})$$

for any nef divisor B. With Theorem 2.16, this implies

$$\widehat{\mathrm{vol}}(B^{n-1}) \leq \mathrm{mob}(B^{n-1}).$$

In general, for a big curve class α we have

$$\mathrm{mob}(\alpha) \geq \sup_{\substack{B \text{ nef},\\ \alpha\succeq B^{n-1}}} \mathrm{mob}(B^{n-1}) \geq \sup_{\substack{B \text{ nef},\\ \alpha\succeq B^{n-1}}} B^n = \widehat{\mathrm{vol}}(\alpha).$$

where the last equality again follows from Theorem 2.16. This finishes the proof.

(2) To prove the second part of Theorem 6.1, we need the following result:

Lemma 6.5 [Fulger and Lehmann 2017b, Corollary 6.16]. *Let X be a smooth projective variety of dimension n and let α be a big curve class. Then there is a big movable curve class β satisfying $\beta \preceq \alpha$ such that*

$$\mathrm{mob}(\alpha) = \mathrm{mob}(\beta) = \mathrm{mob}(\phi^*\beta)$$

for any birational map $\phi : Y \to X$ from a smooth variety Y.

We now prove the statement via a sequence of claims.

Claim. *Assume Conjecture 1.4. If β is a movable curve class with $\mathfrak{M}(\beta) > 0$, then for any $\epsilon > 0$ there is a birational map $\phi_\epsilon : Y_\epsilon \to X$ such that*

$$\mathfrak{M}(\beta) - \epsilon \leq \mathrm{mob}(\phi_\epsilon^*\beta) \leq \mathfrak{M}(\beta) + \epsilon.$$

By Theorem 3.14, we may suppose that there is a big divisor L such that $\beta = \langle L^{n-1}\rangle$. Without loss of generality we may assume that L is effective. Fix an ample effective divisor G as in [Fulger and Lehmann 2017b, Proposition 6.24]; the proposition shows that for any sufficiently small ϵ there is a birational morphism $\phi_\epsilon : Y_\epsilon \to X$ and a big and nef divisor A_ϵ on Y_ϵ satisfying

$$A_\epsilon \leq P_\sigma(\phi_\epsilon^* L) \leq A_\epsilon + \epsilon\phi_\epsilon^* G.$$

Note that $\mathrm{vol}(A_\epsilon) \leq \mathrm{vol}(L) \leq \mathrm{vol}(A_\epsilon + \epsilon\phi_\epsilon^* G)$. Furthermore, we have

$$\mathrm{vol}(A_\epsilon + \epsilon\phi_\epsilon^* G) \leq \mathrm{vol}(\phi_{\epsilon *}A_\epsilon + \epsilon G) \leq \mathrm{vol}(L + \epsilon G).$$

Applying [Fulger and Lehmann 2017b, Lemma 6.21] and the invariance of the positive product under passing to positive parts, we have

$$A_\epsilon^{n-1} \preceq \phi_\epsilon^*\beta \preceq (A_\epsilon + \epsilon\phi_\epsilon^* G)^{n-1}.$$

Applying Conjecture 1.4 (which is only stated for ample divisors but applies to big and nef divisors by continuity of mob), we find

$$\mathrm{vol}(A_\epsilon) = \mathrm{mob}(A_\epsilon^{n-1}) \leq \mathrm{mob}(\phi_\epsilon^*\beta) \leq \mathrm{mob}((A_\epsilon + \epsilon\phi_\epsilon^*(G))^{n-1}) = \mathrm{vol}(A_\epsilon + \epsilon\phi_\epsilon^* G).$$

As ϵ shrinks the two outer terms approach $\mathrm{vol}(L) = \mathfrak{M}(\beta)$.

Claim. *Assume Conjecture 1.4. If a big movable curve class β satisfies $\mathrm{mob}(\beta) = \mathrm{mob}(\phi^*\beta)$ for every birational ϕ then we must have $\beta \in \mathrm{CI}_1(X)$.*

When $\mathfrak{M}(\beta) > 0$, by the previous claim we see from taking a limit that $\mathrm{mob}(\beta) = \mathfrak{M}(\beta)$. By Theorem 6.1(1) and Theorem 5.1 we get

$$\widehat{\mathrm{vol}}(\beta) \leq \mathfrak{M}(\beta) \leq \widehat{\mathrm{vol}}(\beta)$$

and Theorem 5.1 implies the result. When $\mathfrak{M}(\beta) = 0$, fix a class ξ in the interior of the movable cone and consider $\beta + \delta\xi$ for $\delta > 0$. By the previous claim, for any $\epsilon > 0$ we can find a sufficiently small δ and a birational map $\phi_\epsilon : Y_\epsilon \to X$ such that $\mathrm{mob}(\phi_\epsilon^*(\beta + \delta\xi)) < \epsilon$. We also have $\mathrm{mob}(\phi_\epsilon^*\beta) \leq \mathrm{mob}(\phi_\epsilon^*(\beta + \delta\xi))$ since the pullback of the nef curve class $\delta\xi$ is pseudoeffective. By the assumption on the birational invariance of $\mathrm{mob}(\beta)$, we can take a limit to obtain $\mathrm{mob}(\beta) = 0$, a contradiction to the bigness of β.

To finish the proof, recall that Lemma 6.5 implies that the mobility of α must coincide with the mobility of a movable class β lying below α and satisfying $\mathrm{mob}(\pi^*\beta) = \mathrm{mob}(\beta)$ for any birational map π. Thus we have shown

$$\mathrm{mob}(\alpha) = \sup_{\substack{B \text{ nef},\\ \alpha \succeq B^{n-1}}} \mathrm{mob}(B^{n-1}).$$

By Conjecture 1.4 again, we obtain

$$\mathrm{mob}(\alpha) = \sup_{\substack{B \text{ nef},\\ \alpha \succeq B^{n-1}}} B^n.$$

But the right-hand side agrees with $\widehat{\mathrm{vol}}(\alpha)$ by Theorem 2.16. This proves the equality $\mathrm{mob}(\alpha) = \widehat{\mathrm{vol}}(\alpha)$ under the Conjecture 1.4.

(3) We now prove the equality $\widehat{\mathrm{vol}} = \mathrm{wmob}$. The key advantage is that the analogue of Conjecture 1.4 is known for the weighted mobility: Example 8.19 of [Lehmann 2016] shows that for any big and nef divisor B we have $\mathrm{wmob}(B^{n-1}) = B^n$.

We first prove the inequality $\widehat{\mathrm{vol}} \geq \mathrm{wmob}$. The argument is essentially identical to the upper bound in Theorem 6.1(1); by continuity and homogeneity it suffices to prove it for classes in $N_1(X)_{\mathbb{Z}}$. Choose a positive integer μ and a family of curves of class $\mu m\alpha$ achieving $\mathrm{wmc}(m\alpha)$. By splitting up into components and applying Proposition 6.2 with equal weight μ at every point we see that for any component U_i with class β_i we have

$$\widehat{\mathrm{vol}}(\beta_i) \geq \mu^{n/(n-1)}(\mathrm{wmc}(U_i) - 1)$$

Arguing as in Theorem 6.1(1), we see that for any fixed ample Cartier divisor A we have

$$\widehat{\mathrm{vol}}(m\mu\alpha) \geq \mu^{n/(n-1)}(\mathrm{wmc}(m\alpha) - mA \cdot \alpha).$$

Rescaling by μ and taking a limit proves the statement.

We next prove the inequality $\widehat{\mathrm{vol}} \leq \mathrm{wmob}$. Again, the argument is identical to the lower bound in Theorem 6.1(1). It is clear that the weighted mobility can only increase upon adding an effective class. Using continuity and homogeneity, the same is true for any pseudoeffective class. Thus we have

$$\mathrm{wmob}(\alpha) \geq \sup_{\substack{B \text{ nef}, \\ \alpha \succeq B^{n-1}}} \mathrm{wmob}(B^{n-1}) = \sup_{\substack{B \text{ nef}, \\ \alpha \succeq B^{n-1}}} B^n = \widehat{\mathrm{vol}}(\alpha).$$

where the second equality follows from [Lehmann 2016, Example 8.19]. This finishes the proof of the equality $\widehat{\mathrm{vol}} = \mathrm{wmob}$. □

Remark 6.6. We expect Theorem 6.1 to also hold over any algebraically closed field, but we have not thoroughly checked the results on asymptotic multiplier ideals used in the proof of [Fulger and Lehmann 2017b, Proposition 6.24].

Theorem 6.1 yields two interesting consequences:

- The theorem indicates (loosely speaking) that if the mobility count of complete intersection classes is optimized by complete intersection curves, then the mobility count of *any* curve class is optimized by complete intersection curves lying below the class.

  This result is very surprising: it indicates that the "positivity" of a curve class is coming from ample divisors in a strong sense. For example, suppose that X and X' are isomorphic in codimension 1. If we take a complete intersection class α on X, we expect that complete intersections of ample divisors maximize the mobility count. However, the strict transform of these curves on X' should not maximize the mobility count. Instead, if we deform these curves so that they break off a piece contained in the exceptional locus, the part left over will lie in a family which deforms more than the original.
- The theorem suggests that the Zariski decomposition constructed in [Fulger and Lehmann 2017b] for curves is not optimal: instead of defining a positive part in the movable cone, if Conjecture 1.4 is true we should instead define a positive part in the complete intersection cone. It would be interesting to see an analogous improvement for higher dimension cycles.

Asymptotic point counts and $\mathfrak{M}$. Finally, we show that $\mathfrak{M}$ can be given an enumerative interpretation.

Definition 6.7. Let $p : U \to W$ be a family of curves on X with morphism $s : U \to X$. We say that U is strictly movable if:

(1) For each component U_i of U, the morphism $s|_{U_i}$ is dominant.

(2) For each component U_i of U, the morphism $p|_{U_i}$ has generically irreducible fibers.

We then define $\mathrm{mob}_{\mathrm{mov}}$ and $\mathrm{wmob}_{\mathrm{mov}}$ exactly analogously to mob and wmob, except that we only allow contributions of strictly movable families of curves. Note that $\mathrm{mob}_{\mathrm{mov}}$ and $\mathrm{wmob}_{\mathrm{mov}}$ vanish outside of $\mathrm{Mov}_1(X)$ since these classes are not represented by a sum of irreducible curves which deform to dominate X. Arguing just as in [Lehmann 2016, Section 5], one sees that $\mathrm{mob}_{\mathrm{mov}}$ and $\mathrm{wmob}_{\mathrm{mov}}$ are homogeneous of weight $n/(n-1)$, and are continuous in the interior of $\mathrm{Mov}_1(X)$.

Lemma 6.8. *Let $\phi : Y \to X$ be a birational morphism of smooth projective varieties. Let $p : U \to W$ be a family of irreducible curves admitting a dominant map $s : U \to X$. Let U_Y be the family of curves defined by strict transforms. Letting α, α_Y denote respectively the classes of the families on X, Y, we have that $\phi^*\alpha - \alpha_Y$ is the class of an effective $\mathbb{R}$-curve.*

Proof. Since α_Y is the class of a family of irreducible curves which dominates Y, it has nonnegative intersection against every effective divisor. Arguing as in the negativity of contraction lemma, we can find a basis $\{e_i\}$ of $\ker(\phi_* : N_1(Y) \to N_1(X))$ consisting of effective curves and a basis $\{f_j\}$ of $\ker(\phi_* : N^1(Y) \to N^1(X))$ consisting of effective divisors such that the intersection matrix is negative definite and the only negative entries are on the diagonal. Just as in [Bauer et al. 2012, Lemma 4.1], this shows that

$$\alpha_Y = \phi^*\phi_*\alpha_Y - \beta = \phi^*\alpha - \beta$$

for some effective curve class β supported on the exceptional divisors. □

Theorem 6.9. *Let X be a smooth projective variety of dimension n and let $\alpha \in \mathrm{Mov}_1(X)^\circ$. Then*:

(1) $\mathfrak{M}(\alpha) = \mathrm{wmob}_{\mathrm{mov}}(\alpha)$.

(2) *Assume Conjecture 1.4. Then $\mathfrak{M}(\alpha) = \mathrm{mob}_{\mathrm{mov}}(\alpha)$.*

Proof. (1) Suppose that $\phi : Y \to X$ is a birational model of X and that A is an ample Cartier divisor on X. By pushing-forward complete intersection families, we see that $\mathrm{wmob}_{\mathrm{mov}}(\phi_* A^{n-1}) \geq A^n$. By continuity we obtain the inequality $\mathfrak{M}(\alpha) \leq \mathrm{wmob}_{\mathrm{mov}}(\alpha)$ for any $\alpha \in \mathrm{Mov}_1(X)^\circ$.

To see the reverse inequality, by continuity and homogeneity it suffices to consider the case when $\alpha \in \mathrm{Mov}_1(X)^\circ_{\mathbb{Z}}$. Choose a positive integer μ and a strictly movable family of curves U of class $\mu m\alpha$ achieving $\mathrm{wmc}_{\mathrm{mov}}(m\alpha)$. Let $\phi : Y \to X$ be a birational model and let U_Y denote the strict transform class on Y with numerical class α'. By arguing as in the proof of Theorem 6.1, we find that

$$\mathfrak{M}(\alpha') \geq \mu^{n/(n-1)}(\mathrm{wmc}_{\mathrm{mov}}(m\alpha) - mA \cdot \alpha).$$

Furthermore by Lemma 6.8 we have $\widehat{\mathrm{vol}}(m\mu\phi^*\alpha) \geq \widehat{\mathrm{vol}}(\alpha')$. Dividing by $m^{n/(n-1)}$ and taking a limit as m increases, we see that $\mathfrak{M}(\alpha) \geq \mathrm{wmob}_{\mathrm{mov}}(\alpha)$.

(2) The proof of $\mathfrak{M}(\alpha) \leq \mathrm{mob}_{\mathrm{mov}}(\alpha)$ is the same as in (1). Conversely, suppose that U is a strictly movable family of curves achieving $\mathrm{mc}_{\mathrm{mov}}(m\alpha)$. Let $\phi : Y \to X$ be a birational morphism of smooth varieties; by combining Lemma 6.8 with [Fulger and Lehmann 2017b, Section 4], we see that $\mathrm{mc}_{\mathrm{mov}}(m\alpha) \leq \mathrm{mc}_{\mathcal{K}}(m\phi^*\alpha)$, where $\mathcal{K}$ is a cone chosen as in [Fulger and Lehmann 2017b, Definition 4.8] and includes a fixed effective basis of the kernel of $\phi_* : N_1(Y) \to N_1(X)$ chosen as in Lemma 6.8. Taking limits, we see that $\mathrm{mob}_{\mathrm{mov}}(\alpha) \leq \mathrm{mob}(\phi^*\alpha)$ for any birational map ϕ.

Choose a sequence of birational maps $\phi_i : Y_i \to X$ as in the proof of Proposition 5.2 so that $\widehat{\mathrm{vol}}(\phi_i^*\alpha)$ limits to $\mathfrak{M}(\alpha)$. By taking a limit over i and applying Theorem 6.1(2) we finish the proof. □

Acknowledgements

We thank M. Jonsson for his many helpful comments. Some of the material on toric varieties was worked out in a conversation with J. Huh, and we are very grateful for his help. Lehmann would like to thank C. Araujo, M. Fulger, D. Greb, R. Lazarsfeld, S. Payne, D. Treumann, and D. Yang for helpful conversations. Xiao would like to thank his supervisor J.-P. Demailly for suggesting an intersection-theoretic approach to study volume functions, and thank S. Boucksom and W. Ou for helpful conversations, and was supported in part by CSC fellowship and Institut Fourier when this work was prepared. We also would like to thank the referee for helpful comments.

References

[Bauer et al. 2012] T. Bauer, M. Caibăr, and G. Kennedy, "Zariski decomposition: a new (old) chapter of linear algebra", *Amer. Math. Monthly* **119**:1 (2012), 25–41. MR Zbl

[Boucksom 2002a] S. Boucksom, *Cônes positifs des variétés complexes compactes*, Ph.D. thesis, Université Joseph-Fourier-Grenoble I, 2002.

[Boucksom 2002b] S. Boucksom, "On the volume of a line bundle", *Internat. J. Math.* **13**:10 (2002), 1043–1063. MR Zbl

[Boucksom 2004] S. Boucksom, "Divisorial Zariski decompositions on compact complex manifolds", *Ann. Sci. École Norm. Sup.* (4) **37**:1 (2004), 45–76. MR Zbl

[Boucksom et al. 2009] S. Boucksom, C. Favre, and M. Jonsson, "Differentiability of volumes of divisors and a problem of Teissier", *J. Algebraic Geom.* **18**:2 (2009), 279–308. MR Zbl

[Boucksom et al. 2010] S. Boucksom, P. Eyssidieux, V. Guedj, and A. Zeriahi, "Monge–Ampère equations in big cohomology classes", *Acta Math.* **205**:2 (2010), 199–262. MR Zbl

[Boucksom et al. 2013] S. Boucksom, J.-P. Demailly, M. Păun, and T. Peternell, "The pseudo-effective cone of a compact Kähler manifold and varieties of negative Kodaira dimension", *J. Algebraic Geom.* **22**:2 (2013), 201–248. MR Zbl

[Campana 1992] F. Campana, "Connexité rationnelle des variétés de Fano", *Ann. Sci. École Norm. Sup.* (4) **25**:5 (1992), 539–545. MR Zbl

[Cascini et al. 2014] P. Cascini, C. Hacon, M. Mustaţă, and K. Schwede, "On the numerical dimension of pseudo-effective divisors in positive characteristic", *Amer. J. Math.* **136**:6 (2014), 1609–1628. MR Zbl

[Cutkosky 2015] S. D. Cutkosky, "Teissier's problem on inequalities of nef divisors", *J. Algebra Appl.* **14**:9 (2015), art. id. 1540002, 37 pp. MR Zbl

[Debarre et al. 2011] O. Debarre, L. Ein, R. Lazarsfeld, and C. Voisin, "Pseudoeffective and nef classes on abelian varieties", *Compos. Math.* **147**:6 (2011), 1793–1818. MR Zbl

[Demailly 2012] J.-P. Demailly, *Complex analytic and differential geometry*, 2012. Zbl

[Demailly et al. 2014] J.-P. Demailly, S. a. Dinew, V. Guedj, H. H. Pham, S. a. Koł odziej, and A. Zeriahi, "Hölder continuous solutions to Monge–Ampère equations", *J. Eur. Math. Soc. (JEMS)* **16**:4 (2014), 619–647. MR Zbl

[Fulger 2011] M. Fulger, "The cones of effective cycles on projective bundles over curves", *Math. Z.* **269**:1-2 (2011), 449–459. MR Zbl

[Fulger and Lehmann 2017a] M. Fulger and B. Lehmann, "Positive cones of dual cycle classes", *Algebr. Geom.* **4**:1 (2017), 1–28. MR Zbl

[Fulger and Lehmann 2017b] M. Fulger and B. Lehmann, "Zariski decompositions of numerical cycle classes", *J. Algebraic Geom.* **26**:1 (2017), 43–106. MR Zbl

[Fulton 1984] W. Fulton, *Intersection theory*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)] **2**, Springer, 1984. MR Zbl

[Greb et al. 2016a] D. Greb, S. Kebekus, and T. Peternell, "Movable curves and semistable sheaves", *Int. Math. Res. Not.* **2016**:2 (2016), 536–570. MR Zbl

[Greb et al. 2016b] D. Greb, J. Ross, and M. Toma, "A master space for moduli spaces of Gieseker-stable sheaves", 2016. arXiv

[Greb et al. 2016c] D. Greb, J. Ross, and M. Toma, "Moduli of vector bundles on higher-dimensional base manifolds—construction and variation", *Internat. J. Math.* **27**:6 (2016), art. id. 1650054, 27 pp. MR Zbl

[Greb et al. 2016d] D. Greb, J. Ross, and M. Toma, "Variation of Gieseker moduli spaces via quiver GIT", *Geom. Topol.* **20**:3 (2016), 1539–1610. MR Zbl

[Greb et al. 2019] D. Greb, J. Ross, and M. Toma, "Semi-continuity of stability for sheaves and variation of Gieseker moduli spaces", *J. Reine Angew. Math.* **749** (2019), 227–265. MR Zbl

[Hu and Keel 2000] Y. Hu and S. Keel, "Mori dream spaces and GIT", *Michigan Math. J.* **48** (2000), 331–348. MR Zbl

[Klain 2004] D. A. Klain, "The Minkowski problem for polytopes", *Adv. Math.* **185**:2 (2004), 270–288. MR Zbl

[Kollár 1996] J. Kollár, *Rational curves on algebraic varieties*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics] **32**, Springer, 1996. MR

[Kollár et al. 1992] J. Kollár, Y. Miyaoka, and S. Mori, "Rational connectedness and boundedness of Fano manifolds", *J. Differential Geom.* **36**:3 (1992), 765–779. MR Zbl

[Küronya 2006] A. Küronya, "Asymptotic cohomological functions on projective varieties", *Amer. J. Math.* **128**:6 (2006), 1475–1519. MR Zbl

[Küronya and Maclean 2013] A. Küronya and C. Maclean, "Zariski decomposition of b-divisors", *Math. Z.* **273**:1-2 (2013), 427–436. MR Zbl

[Lazarsfeld 2004] R. Lazarsfeld, *Positivity in algebraic geometry, I*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics] **48**, Springer, 2004. MR Zbl

[Lazarsfeld and Mustaţă 2009] R. Lazarsfeld and M. Mustaţă, "Convex bodies associated to linear series", *Ann. Sci. Éc. Norm. Supér.* (4) **42**:5 (2009), 783–835. MR Zbl

[Lehmann 2016] B. Lehmann, "Volume-type functions for numerical cycle classes", *Duke Math. J.* **165**:16 (2016), 3147–3187. MR Zbl

[Lehmann and Xiao 2016] B. Lehmann and J. Xiao, "Convexity and Zariski decomposition structure", *Geom. Funct. Anal.* **26**:4 (2016), 1135–1189. MR Zbl

[Lehmann and Xiao 2017] B. Lehmann and J. Xiao, "Correspondences between convex geometry and complex geometry", *Épijournal Geom. Algébrique* **1** (2017), Art. id. 6, 29 pp. MR Zbl

[Mustaţă 2013] M. Mustaţă, "The non-nef locus in positive characteristic", pp. 535–551 in *A celebration of algebraic geometry*, edited by B. Hassett et al., Clay Math. Proc. **18**, Amer. Math. Soc., Providence, RI, 2013. MR Zbl

[Nakayama 2004] N. Nakayama, *Zariski-decomposition and abundance*, MSJ Memoirs **14**, Mathematical Society of Japan, Tokyo, 2004. MR Zbl

[Neumann 2010] S. Neumann, *A decomposition of the Moving cone of a projective manifold according to the Harder–Narasimhan filtration of the tangent bundle*, Ph.D. thesis, Mathematisches Institut der Universität, 2010. Zbl

[Nyström and Boucksom 2016] D. W. Nyström and S. Boucksom, "Duality between the pseudoeffective and the movable cone on a projective manifold", 2016. arXiv

[Perrin 1987] D. Perrin, "Courbes passant par m points généraux de $\mathbf{P}^3$", *Mém. Soc. Math. France* (*N.S.*) 28-29 (1987), 138. MR Zbl

[Rockafellar 1970] R. T. Rockafellar, *Convex analysis*, Princeton Mathematical Series, No. 28 **28**, Princeton University Press, 1970. MR Zbl

[Takagi 2007] S. Takagi, "Fujita's approximation theorem in positive characteristics", *J. Math. Kyoto Univ.* **47**:1 (2007), 179–202. MR Zbl

[Trapani 1995] S. Trapani, "Numerical criteria for the positivity of the difference of ample divisors", *Math. Z.* **219**:3 (1995), 387–401. MR Zbl

[Xiao 2017] J. Xiao, "Characterizing volume via cone duality", *Math. Ann.* **369**:3-4 (2017), 1527–1555. MR Zbl

[Xiao 2018] J. Xiao, "Movable intersection and bigness criterion", *Universitatis Iagellonicae Acta Mathematica* **2018**:55 (2018), 53–64.

[Zariski 1962] O. Zariski, "The theorem of Riemann–Roch for high multiples of an effective divisor on an algebraic surface", *Ann. of Math.* (2) **76** (1962), 560–615. MR Zbl

Communicated by Ravi Vakil
Received 2018-03-02 Revised 2019-02-20 Accepted 2019-04-08

lehmannb@bc.edu — *Department of Mathematics, Boston College, Chestnut Hill, MA, United States*

jianxiao@tsinghua.edu.cn — *Department of Mathematical Sciences and Yau Mathematical Sciences Center, Tsinghua University, Beijing, China*

ALGEBRA AND NUMBER THEORY 13:6 (2019)
dx.doi.org/10.2140/ant.2019.13.1281

The congruence topology, Grothendieck duality and thin groups

Alexander Lubotzky and Tyakal Nanjundiah Venkataramana

This paper answers a question raised by Grothendieck in 1970 on the "Grothendieck closure" of an integral linear group and proves a conjecture of the first author made in 1980. This is done by a detailed study of the congruence topology of arithmetic groups, obtaining along the way, an arithmetic analogue of a classical result of Chevalley for complex algebraic groups. As an application we also deduce a group theoretic characterization of thin subgroups of arithmetic groups.

Introduction

If $\varphi\colon G_1 \to G_2$ is a polynomial map between two complex varieties, then in general the image of a Zariski closed subset of G_1 is not necessarily closed in G_2. But here is a classical result:

Theorem (Chevalley). *If φ is a polynomial homomorphism between two complex algebraic groups then $\varphi(H)$ is closed in G_2 for every closed subgroup H of G_1.*

There is an arithmetic analogue of this issue: Let G be a $\mathbb{Q}$-algebraic group, let $\mathbb{A}_f = \Pi^*_{p \text{ prime}} \mathbb{Q}_p$ be the ring of finite adeles over $\mathbb{Q}$. The topology of $G(\mathbb{A}_f)$ induces the congruence topology on $G(\mathbb{Q})$. If K is a compact open subgroup of $G(\mathbb{A}_f)$ then $\Gamma = K \cap G(\mathbb{Q})$ is called a congruence subgroup of $G(\mathbb{Q})$. This defines the congruence topology on $G(\mathbb{Q})$ and on all its subgroups. A subgroup of $G(\mathbb{Q})$ which is closed in this topology is called congruence closed. A subgroup Δ of G commensurable to Γ is called an arithmetic group.

Now, if $\varphi\colon G_1 \to G_2$ is a $\mathbb{Q}$-morphism between two $\mathbb{Q}$-groups, which is a surjective homomorphism (as $\mathbb{C}$-algebraic groups) then the image of an arithmetic subgroup Δ of G_1 is an arithmetic subgroup of G_2 [Platonov and Rapinchuk 1994, Theorem 4.1, page 74], but the image of a congruence subgroup is not necessarily a congruence subgroup. It is well known that $\mathrm{SL}_n(\mathbb{Z})$ has congruence subgroups whose images under the adjoint map $\mathrm{SL}_n(\mathbb{Z}) \to \mathrm{PSL}_n(\mathbb{Z}) \hookrightarrow \mathrm{Aut}(M_n(\mathbb{Z}))$ are not congruence subgroups (see [Serre 1968] and Proposition 2.1 below for an exposition and explanation). So, the direct analogue of Chevalley's theorem does not hold. Still, in this case, if Γ is a congruence subgroup of $\mathrm{SL}_n(\mathbb{Z})$, then $\varphi(\Gamma)$ is a normal subgroup of $\overline{\varphi(\Gamma)}$, the (congruence) closure of $\varphi(\Gamma)$ in $\mathrm{PSL}_n(\mathbb{Z})$, and the quotient is a finite abelian group. Our first technical result says that the general case is similar. It is especially important

MSC2010: primary 11E57; secondary 20G30.
Keywords: congruence subgroup, thin groups.

for us that when G_2 is simply connected, the image of a congruence subgroup of G_1 is a congruence subgroup in G_2 (see Proposition 0.1(ii) below).

Before stating the result, we give the following definition and set some notations for the rest of the paper.

Let G be a linear algebraic group over $\mathbb{C}$, G^0 its connected component, and $R = R(G)$ its solvable radical, i.e., the largest connected normal solvable subgroup of G. We say that G is *essentially simply connected* if $G_{ss} := G^0/R$ is simply connected.

Given a subgroup Γ of GL_n, we will throughout the paper denote by Γ^0 the intersection of Γ with G^0, where G^0 is the connected component of G, the Zariski closure of Γ. Therefore, Γ^0 is always a finite index normal subgroup of Γ.

The notion "essentially simply connected" will play an important role in this paper due to the following proposition, which can be considered as the arithmetic analogue of Chevalley's result above.

Proposition 0.1. (i) *If $\varphi\colon G_1 \to G_2$ is a surjective $\mathbb{Q}$-morphism of algebraic $\mathbb{Q}$-groups, then for every congruence closed subgroup Γ of $G_1(\mathbb{Q})$, the image $\varphi(\Gamma^0)$ is normal in its congruence closure $\overline{\varphi(\Gamma^0)}$ and $\overline{\varphi(\Gamma^0)}/\varphi(\Gamma^0)$ is a finite abelian group.*

(ii) *If G_2 is essentially simply connected and Γ a congruence subgroup of G_1 then $\overline{\varphi(\Gamma)} = \varphi(\Gamma)$, i.e., the image of a congruence subgroup is congruence closed.*

This analogue of Chevalley's theorem and a result of [Nori 1987; Weisfeiler 1984] enable us to prove:

Proposition 0.2. *If $\Gamma_1 \le \mathrm{GL}_n(\mathbb{Z})$ is a congruence closed subgroup (i.e., closed in the congruence topology) with Zariski closure G, then there exists a congruence subgroup Γ of G, such that $[\Gamma, \Gamma] \le \Gamma_1^0 \le \Gamma$. If G is essentially simply connected then the image of Γ_1 in $G/R(G)$ is actually a congruence subgroup.*

We apply Proposition 0.1(ii) in two directions:

(A) Grothendieck–Tannaka duality for discrete groups, and

(B) a group theoretic characterization of thin subgroups of arithmetic groups.

Grothendieck closure. Grothendieck [1970] was interested in the following question:

Question 0.3. Assume $\varphi\colon \Gamma_1 \to \Gamma_2$ is a homomorphism between two finitely generated residually finite groups inducing an isomorphism $\hat{\varphi}\colon \hat{\Gamma}_1 \to \hat{\Gamma}_2$ between their profinite completions. Is φ already an isomorphism?

To tackle Question 0.3, he introduced the following notion. Given a finitely generated group Γ and a commutative ring A with identity, let $\mathrm{Cl}_A(\Gamma)$ be the group of all automorphisms of the forgetful functor from the category $\mathrm{Mod}_A(\Gamma)$ of all finitely generated A-modules with Γ action to $\mathrm{Mod}_A(\{1\})$, preserving tensor product. Recall that α is such an automorphism means that for every finitely generated A-module L with Γ action $\rho_L\colon \Gamma \to \mathrm{Aut}_A(L)$, we are given $\alpha_L \in \mathrm{Aut}_A(L)$ such that if $\varphi\colon L_1 \to L_2$ is an $A[\Gamma]$-morphism between such modules then $\alpha_{L_2} \circ \varphi = \varphi \circ \alpha_{L_1}$. In particular, every $\tau \in \Gamma$ defines such α, by $\alpha_L = \rho_L(\tau)$. This gives a natural map from Γ to $\mathrm{Cl}_A(\Gamma)$.

Grothendieck's strategy was the following: he showed that, under the conditions of Question 0.3, φ induces an isomorphism from $\mathrm{Mod}_A(\Gamma_2)$ to $\mathrm{Mod}_A(\Gamma_1)$, and hence also between $\mathrm{Cl}_A(\Gamma_1)$ and $\mathrm{Cl}_A(\Gamma_2)$. He then asked:

Question 0.4. Is the natural map $\Gamma \hookrightarrow \mathrm{Cl}_{\mathbb{Z}}(\Gamma)$ an isomorphism for a finitely generated residually finite group?

An affirmative answer to Question 0.4 would imply an affirmative answer to Question 0.3. Grothendieck then showed that arithmetic groups with the (strict) congruence subgroup property do indeed satisfy $\mathrm{Cl}_{\mathbb{Z}}(\Gamma) \simeq \Gamma$. For a general survey on the congruence subgroup problem see [Raghunathan 1991].

Question 0.4 basically asks whether Γ can be recovered from its category of representations. The first author [Lubotzky 1980] phrased this question in the framework of Tannaka duality, which asks a similar question for compact Lie groups. He also gave a more concrete description of $\mathrm{Cl}_{\mathbb{Z}}(\Gamma)$:

$$\mathrm{Cl}_{\mathbb{Z}}(\Gamma) = \{g \in \hat{\Gamma} \mid \hat{\rho}(g)(V) = V,\ \forall (\rho, V) \in \mathrm{Mod}_{\mathbb{Z}}(\Gamma)\}. \tag{0-1}$$

Here $\hat{\rho}$ is the continuous extension $\hat{\rho}\colon \hat{\Gamma} \to \mathrm{Aut}(\hat{V})$ of the original representation $\rho\colon \Gamma \to \mathrm{Aut}(V)$.

However, it is also shown in [Lubotzky 1980], that the answer to Question 0.4 is negative. The counterexamples provided there are the arithmetic groups for which the weak congruence subgroup property holds but not the strict one, i.e., the congruence kernel is finite but nontrivial. It was conjectured in [Lubotzky 1980, Conjecture A, page 184], that for an arithmetic group Γ, $\mathrm{Cl}_{\mathbb{Z}}(\Gamma) = \Gamma$ if and only if Γ has the (strict) congruence subgroup property. The conjecture was left open even for $\Gamma = \mathrm{SL}_2(\mathbb{Z})$.

In the almost 40 years since [Lubotzky 1980] was written various counterexamples were given to Question 0.3 [Platonov and Tavgen 1986; Bass and Lubotzky 2000; Bridson and Grunewald 2004; Pyber 2004] which also give counterexamples to Question 0.4, but it was not even settled whether $\mathrm{Cl}_{\mathbb{Z}}(F) = F$ for finitely generated nonabelian free groups F.

We can now answer this and, in fact, prove the following surprising result, which gives an essentially complete answer to Question 0.4.

Theorem 0.5. *Let Γ be a finitely generated subgroup of $\mathrm{GL}_n(\mathbb{Z})$. Then Γ satisfies Grothendieck–Tannaka duality, i.e., $\mathrm{Cl}_{\mathbb{Z}}(\Gamma) = \Gamma$ if and only if Γ has the congruence subgroup property i.e., for some (and consequently for every) faithful representation $\Gamma \to \mathrm{GL}_m(\mathbb{Z})$ such that the Zariski closure G of Γ is essentially simply connected, every finite index subgroup of Γ is closed in the congruence topology of $\mathrm{GL}_n(\mathbb{Z})$. In such a case, the image of the group Γ in the semisimple (simply connected) quotient G/R is a congruence arithmetic group.*

The theorem is surprising as it shows that the cases proved by Grothendieck himself (which motivated him to suggest that the duality holds in general) are essentially the only cases where this duality holds.

Let us note that the assumption on G is not really restrictive. In Lemma 3.5, we show that for every $\Gamma \leq \mathrm{GL}_n(\mathbb{Z})$ we can find an "over" representation of Γ into $\mathrm{GL}_m(\mathbb{Z})$ (for some m) whose Zariski closure is essentially simply connected.

Theorem 0.5 implies Conjecture A of [Lubotzky 1980].

Corollary 0.6. *If G is a simply connected semisimple $\mathbb{Q}$-algebraic group, and Γ a congruence subgroup of $G(\mathbb{Q})$, then $\mathrm{Cl}_{\mathbb{Z}}(\Gamma)=\Gamma$ if and only if Γ satisfies the (strict) congruence subgroup property.*

In particular:

Corollary 0.7. $\mathrm{Cl}_{\mathbb{Z}}(F)\neq F$ *for every finitely generated free group on at least two generators; furthermore,* $\mathrm{Cl}_{\mathbb{Z}}(\mathrm{SL}_2(\mathbb{Z}))\neq \mathrm{SL}_2(\mathbb{Z})$.

In fact, it will follow from our results that $\mathrm{Cl}_{\mathbb{Z}}(F)$ is uncountable.

Before moving on to the last application, let us say a few words about how Proposition 0.1 helps to prove a result like Theorem 0.5. The description of $\mathrm{Cl}_{\mathbb{Z}}(\Gamma)$ as in (0-1) implies that

$$\mathrm{Cl}_{\mathbb{Z}}(\Gamma)=\varprojlim_{\rho}\overline{\rho(\Gamma)} \tag{0-2}$$

where the limit is over all (ρ, V), where V is a finitely generated abelian group, ρ a representation $\rho\colon \Gamma\to \mathrm{Aut}(V)$ and $\overline{\rho(\Gamma)}=\hat{\rho}(\hat{\Gamma})\cap \mathrm{Aut}(V)\subseteq \mathrm{Aut}(\hat{V})$. This is an inverse limit of countable discrete groups, so one can not say much about it unless the connecting homomorphisms are surjective, which is, in general, not the case. Now, $\overline{\rho(\Gamma)}$ is the congruence closure of $\rho(\Gamma)$ in $\mathrm{Aut}(V)$ and Proposition 0.1 shows that the corresponding maps are "almost" onto, and are even surjective if the modules V are what we call here "simply connected representations", namely those cases where V is torsion free (and hence isomorphic to $\mathbb{Z}^n$ for some n) and the Zariski closure of $\rho(\Gamma)$ in $\mathrm{Aut}(\mathbb{C}\otimes_{\mathbb{Z}} V)=\mathrm{GL}_n(\mathbb{C})$ is essentially simply connected. We show further that the category $\mathrm{Mod}_{\mathbb{Z}}(\Gamma)$ is "saturated" with such modules (see Lemma 3.5) and we deduce that one can compute $\mathrm{Cl}_{\mathbb{Z}}(\Gamma)$ as in (0-1) by considering only simply connected representations. We can then use Proposition 0.1(b), and get a fairly good understanding of $\mathrm{Cl}_{\mathbb{Z}}(\Gamma)$. This enables us to prove Theorem 0.5. In addition, we also deduce:

Corollary 0.8. *If (ρ, V) is a simply connected representation, then the induced map* $\mathrm{Cl}_{\mathbb{Z}}(\Gamma)\to \mathrm{Aut}(V)$ *is onto* $\mathrm{Cl}_{\rho}(\Gamma):=\overline{\rho(\Gamma)}$ *— the congruence closure of* Γ.

From Corollary 0.8 we can deduce our last application.

Thin groups. In recent years, following [Sarnak 2014] (see also [Kontorovich et al. 2019]), there has been a lot of interest in the distinction between thin subgroups and arithmetic subgroups of algebraic groups. Let us recall:

Definition 0.9. A subgroup $\Gamma\leq \mathrm{GL}_n(\mathbb{Z})$ is called *thin* if it is of infinite index in $G\cap \mathrm{GL}_n(\mathbb{Z})$, when G is its Zariski closure in GL_n. For a general group Γ, we will say that it is a *thin group* (or it *has a thin representation*) if for some n there exists a representation $\rho\colon \Gamma\to \mathrm{GL}_n(\mathbb{Z})$ for which $\rho(\Gamma)$ is thin.

During the last five decades a lot of attention was given to the study of arithmetic groups, with many remarkable results, especially for those of higher rank (see [Margulis 1991; Platonov and Rapinchuk 1994]). Much less is known about thin groups. For example, it is not known if there exists a thin group with property (T). Also, given a subgroup of an arithmetic group (say, given by a set of generators) it is difficult to decide whether it is thin or arithmetic (i.e., of finite or infinite index in its integral Zariski closure).

It is therefore of interest and perhaps even surprising that our results enable us to give a purely group theoretical characterization of thin groups $\Gamma \subset \mathrm{GL}_n(\mathbb{Z})$. Before stating the precise result, we make the topology on $\mathrm{Cl}_{\mathbb{Z}}(\Gamma)$ explicit. If we take the class of simply connected representations (ρ, V) for computing the group $\mathrm{Cl}_{\mathbb{Z}}(\Gamma)$, one can then show that $\mathrm{Cl}_{\mathbb{Z}}(\Gamma)/\Gamma$ is a *closed* subspace of the product $\prod_\rho (\mathrm{Cl}_\rho(\Gamma)/\Gamma)$, where each $\mathrm{Cl}_\rho(\Gamma)/\Gamma$ is given the discrete topology. This is the topology on the quotient space $\mathrm{Cl}_{\mathbb{Z}}(\Gamma)/\Gamma$ in the following theorem. We can now state:

Theorem 0.10. *Let* Γ *be finitely generated* $\mathbb{Z}$*-linear group. Then* Γ *is a thin group if and only if it satisfies* (*at least*) *one of the following conditions*:

(1) Γ *is not* FAb (*namely*, *it does have a finite index subgroup with an infinite abelianization*).

(2) $\mathrm{Cl}_{\mathbb{Z}}(\Gamma)/\Gamma$ *is not compact.*

Warning. There are groups Γ which can be realized both as arithmetic groups as well as thin groups. For example, the free group is an arithmetic subgroup of $\mathrm{SL}_2(\mathbb{Z})$, but at the same time a thin subgroup of every semisimple group, by a well-known result of Tits [1972]. In our terminology this is a thin group.

Remark 0.11. In the present paper, we have concentrated only on $\mathbb{Z}$ modules which are Γ modules, and we have assumed that $\Gamma \subset \mathrm{SL}_n(\mathbb{Z})$ for some n, so Γ is either a $\mathbb{Q}$-arithmetic group or a thin subgroup of such. Note that if $\mathbb{Z}$ is replaced by O — the ring of integers in a number field K — then by restriction of scalars every arithmetic subgroup of a K-algebraic group is commensurable to such one over $\mathbb{Q}$, so there is no loss of generality here. Moreover, the ring $\mathbb{Z}$ can be replaced by the ring of S-integers O_S, S is a finite set of places including all the archimedean ones and O_S is the subring of K consisting of elements x in K such that for every (finite) place $v \notin S$ of K, x lies in the maximal compact subring O_v of K_v (K_v is the completion of K at the place v). To be precise, one can consider finitely generated O_S modules (in place of $\mathbb{Z}$ modules) which are Γ modules, and one assumes that $\Gamma \subset \mathrm{SL}_n(O_S)$ i.e., Γ is a subgroup of an S-arithmetic group. One can then talk of the O_S-closure $\mathrm{Cl}_{O_S}(\Gamma)$. The statements and proofs are almost identical, if notationally more tedious and hence we do not wish to pursue this further.

Note however, that one should not "mix between rings", for example, if $\Gamma = \mathrm{SL}_n\big(\mathbb{Z}\big[\frac{1}{p}\big]\big)$, $n \geq 2$, p a prime, then $\mathrm{Cl}_{\mathbb{Z}[1/p]}(\Gamma) = \Gamma$ as Γ satisfies the congruence subgroup property (CSP). But

$$\mathrm{Cl}_{\mathbb{Z}}(\Gamma) = \hat{\Gamma} = \mathrm{SL}_n\big(\widehat{\mathbb{Z}\big[\tfrac{1}{p}\big]}\big) = \prod_{q \neq p} \mathrm{SL}_n(\mathbb{Z}_q),$$

since every representation of Γ on a finitely generated $\mathbb{Z}$-module factors through a finite quotient.

Notation 0.12. Throughout the paper, if W is a finitely generate $\mathbb{Z}$ module, we denote by $W_{\mathbb{Q}}$ and $W_{\mathbb{C}}$ the $\mathbb{Q}$ vector space $W \otimes_{\mathbb{Z}} \mathbb{Q}$ and the $\mathbb{C}$ vector space $W \otimes_{\mathbb{Z}} \mathbb{C}$, respectively. If $\Gamma \subset \mathrm{GL}(W)$, and G denotes the Zariski closure of Γ in $\mathrm{GL}(W_{\mathbb{C}})$, then G acts on $W_{\mathbb{C}}$. The Zariski closure G is an algebraic group defined over $\mathbb{Q}$ with respect to the $\mathbb{Q}$ structure on $W_{\mathbb{C}} = W_{\mathbb{Q}} \otimes_{\mathbb{Q}} \mathbb{C}$. It is well-known that $G(\mathbb{Q})$ is Zariski dense in G and $G(\mathbb{Q})$ acts on $W_{\mathbb{Q}}$. The action of G on $W_{\mathbb{C}}$ is completely determined by the action of $G(\mathbb{Q})$ on $W_{\mathbb{Q}}$. For these reasons, we do not distinguish between G acting on $W_{\mathbb{C}}$ and $G(\mathbb{Q})$ (by a mild abuse of notation, denoted G) acting on $W_{\mathbb{Q}}$.

1. Preliminaries on algebraic groups over $\mathbb{Q}$

We recall the definition of an essentially simply connected group.

Definition 1.1. Let G be a linear algebraic group over $\mathbb{C}$ with maximal connected normal solvable subgroup R (i.e., the radical of G) and identity component G^0. We say that G is *essentially simply connected* if the semisimple part $G^0/R = H$ is a simply connected.

Note that G is essentially simply connected if and only if, the quotient G^0/U of the group G^0 by its unipotent radical U is a product $H_{\mathrm{ss}} \times S$ with H_{ss} simply connected and semisimple, and S is a torus.

For example, a semisimple connected group is essentially simply connected if and only if it is simply connected. The group $\mathbb{G}_m \times \mathrm{SL}_n$ is essentially simply connected; however, the radical of the group GL_n is the group R of scalars and $\mathrm{GL}_n / R = \mathrm{SL}_n$ /center, so GL_n is *not* essentially simply connected. We will show later (Lemma 1.3(iii)) that every group has a finite cover which is essentially simply connected.

Lemma 1.2. *Suppose $G \subset G_1 \times G_2$ is a subgroup of a product of two essentially simply connected linear algebraic groups G_1, G_2 over $\mathbb{C}$; suppose that the projection π_i of G to G_i is surjective for $i = 1, 2$. Then G is also essentially simply connected.*

Proof. Assume, as we may, that G is connected. Let R be the radical of G. The projection of R to G_i is normal in G_i since $\pi_i\colon G \to G_i$ is surjective. Moreover, $G_i/\pi_i(R)$ is the image of the semisimple group G/R; the latter has a Zariski dense compact subgroup, hence so does $G_i/\pi_i(R)$; therefore, $G_i/\pi_i(R)$ is reductive and is its own commutator. Hence $G_i/\pi_i(R)$ is semisimple and hence $\pi_i(R) = R_i$ where R_i is the radical of G_i. Let $R^* = G \cap (R_1 \times R_2)$. Since $R_1 \times R_2$ is the radical of $G_1 \times G_2$, it follows that R^* is a solvable normal subgroup of G and hence its connected component is contained in R. Since $R \subseteq R_1 \times R_2$, it follows that R is precisely the connected component of the identity of R^*. We then have the inclusion $G/R^* \subset G_1/R_1 \times G_2/R_2$ with projections again being surjective.

By assumption, each $G_i/R_i = H_i$ is semisimple, simply connected. Moreover $G/R^* = H$ where H is connected, semisimple. Thus we have the inclusion $H \subset H_1 \times H_2$. Now, $H \subset H_1 \times H_2$ is such that the projections of H to H_i are surjective, and each H_i is simply connected. Let K be the kernel of the map $H \to H_1$ and K^0 its identity component. Then $H/K^0 \to H_1$ is a surjective map of connected algebraic groups with finite kernel. The simple connectedness of H_1 then implies that $H/K^0 = H_1$ and hence that $K = K^0 \subset \{1\} \times H_2$ is normal in H_2.

Write $H_2 = F_1 \times \cdots \times F_t$ where each F_i is *simple* and simply connected. Now, K being a closed normal subgroup of H_2 must be equal to $\prod_{i \in X} F_i$ for some subset X of $\{1, \ldots, t\}$, and is simply connected. Therefore, $K = K^0$ is simply connected.

From the preceding two paragraphs, we have that both H/K and K are simply connected, and hence so is $H = G/R^*$. Since R is the connected component of R^* and G/R^* is simply connected, it follows that $G/R = G/R^*$ and hence G/R is simply connected. This completes the proof of the lemma. $\square$

Arithmetic groups and congruence subgroups. In the introduction, we defined the notion of arithmetic and congruence subgroup of $G(\mathbb{Q})$ using the adelic language. One can define the notion of arithmetic and congruence groups in more concrete terms as follows. Given a linear algebraic group $G \subset \mathrm{SL}_n$ defined over $\mathbb{Q}$, we will say that a subgroup $\Gamma \subset G(\mathbb{Q})$ is an *arithmetic group* if is commensurable to $G \cap \mathrm{SL}_n(\mathbb{Z}) = G(\mathbb{Z})$; that is, the intersection $\Gamma \cap G(\mathbb{Z})$ has finite index both in Γ and in $G(\mathbb{Z})$. It is well known that the notion of an arithmetic group does not depend on the specific linear embedding $G \subset \mathrm{SL}_n$. As in [Serre 1968], we may define the *arithmetic completion* $\hat{G}$ of $G(\mathbb{Q})$ as the completion of the group $G(\mathbb{Q})$ with respect to the topology on $G(\mathbb{Q})$ as a topological group, obtained by designating arithmetic groups as a fundamental systems of neighborhoods of identity in $G(\mathbb{Q})$.

Given $G \subset \mathrm{SL}_n$ as in the preceding paragraph, we will say that an arithmetic group $\Gamma \subset G(\mathbb{Q})$ is a *congruence subgroup* if there exists an integer $m \geq 2$ such that Γ contains the "principal congruence subgroup" $G(m\mathbb{Z}) = \mathrm{SL}_n(m\mathbb{Z}) \cap G$ where $\mathrm{SL}_n(m\mathbb{Z})$ is the kernel to the residue class map $\mathrm{SL}_n(\mathbb{Z}) \to \mathrm{SL}_n(\mathbb{Z}/m\mathbb{Z})$. We then get the structure of a topological group on the group $G(\mathbb{Q})$ by designating congruence subgroups of $G(\mathbb{Q})$ as a fundamental system of neighborhoods of identity. The completion of $G(\mathbb{Q})$ with respect to this topology, is denoted $\overline{G}$. Again, the notion of a congruence subgroup does not depend on the specific linear Q-embedding $G \to \mathrm{SL}_n$.

Since every congruence subgroup is an arithmetic group, there exists a map from $\pi : \hat{G} \to \overline{G}$ which is easily seen to be surjective, and the kernel $C(G)$ of π is a compact profinite subgroup of $\hat{G}$. This is called the *congruence subgroup kernel*. One says that $G(\mathbb{Q})$ has the *congruence subgroup property* if $C(G)$ is trivial. This is easily seen to be equivalent to the statement that every arithmetic subgroup of $G(\mathbb{Q})$ is a congruence subgroup.

It is known (see page 108, last but one paragraph of [Raghunathan 1976] or [Chahal 1980]) that solvable groups G have the congruence subgroup property.

Moreover, every solvable subgroup of $\mathrm{GL}_n(\mathbb{Z})$ is polycyclic. In such a group, every subgroup is intersection of finite index subgroups. So every solvable subgroup of an arithmetic group is congruence closed. We will use these facts frequently in the sequel.

Another (equivalent) way of viewing the congruence completion is (see [Serre 1968, page 276, Remarque]) as follows: let $\mathbb{A}_f$ be the ring of finite adeles over $\mathbb{Q}$, equipped with the standard adelic topology and let $\mathbb{Z}_f \subset \mathbb{A}_f$ be the closure of $\mathbb{Z}$. Then the group $G(\mathbb{A}_f)$ is also a locally compact group and contains the group $G(\mathbb{Q})$. The congruence completion $\overline{G}$ of $G(\mathbb{Q})$ may be viewed as the closure of $G(\mathbb{Q})$ in $G(\mathbb{A}_f)$.

Lemma 1.3. *Let H and H^* be linear algebraic groups defined over $\mathbb{Q}$.*

(i) *Suppose $\pi : H^* \to H$ is a surjective $\mathbb{Q}$-morphism. Let $(\rho, W_{\mathbb{Q}})$ be a representation of H defined over $\mathbb{Q}$. Then there exists a faithful $\mathbb{Q}$-representation $(\tau, V_{\mathbb{Q}})$ of H^* such that $(\rho \circ \pi, W_{\mathbb{Q}})$ is a subrepresentation of $(\tau, V_{\mathbb{Q}})$.*

(ii) *If $H^* \to H$ is a surjective map defined over $\mathbb{Q}$, then the image of an arithmetic subgroup of H^* under the map $H^* \to H$ is an arithmetic subgroup of H.*

(iii) *If H is connected, then there exists a connected essentially simply connected algebraic group H^* with a surjective $\mathbb{Q}$-defined homomorphism $H^* \to H$ with finite kernel.*

(iv) *If $H^* \to H$ is a surjective homomorphism of algebraic $\mathbb{Q}$-groups which are essentially simply connected, then the image of a congruence subgroup of $H^*(\mathbb{Q})$ is a congruence subgroup of $H(\mathbb{Q})$.*

Proof. Let $\theta\colon H^* \to \mathrm{GL}(E)$ be a faithful representation of the linear algebraic group H^* defined over $\mathbb{Q}$ and $\tau = \rho \oplus \theta$ as H^*-representation. Clearly τ is faithful for H^* and contains ρ. This proves (i).

Part (ii) is the statement of Theorem (4.1) of [Platonov and Rapinchuk 1994].

We now prove (iii). Write $H = RG$ as a product of its radical R and a semisimple group G. Let $H^*_{\mathrm{ss}} \to G$ be the simply connected cover of G. Hence H^*_{ss} acts on R through G, via this covering map. Define $H^* = R \rtimes H^*_{\mathrm{ss}}$ as a semidirect product. Clearly, the map $H^* \to H$ has finite kernel and satisfies the properties of (iii).

To prove (iv), we may assume that H and H^* are connected. If U^* and U are the unipotent radicals of H^* and H, the assumptions of (iv) do not change for the quotient groups H^*/U^* and H/U. Moreover, since H^* is the semidirect product of U^* and H^*/U^* (and similarly for H and U) and the unipotent $\mathbb{Q}$-algebraic group U has the congruence subgroup property, it suffices to prove (iv) when both H^* and H are reductive. By assumption, H^* and H are essentially simply connected; i.e., $H^* = H^*_{\mathrm{ss}} \times S^*$ and $H = H_{\mathrm{ss}} \times S$ where S, S^* are tori and H^*_{ss}, H_{ss} are simply connected semisimple groups. Thus we have connected reductive $\mathbb{Q}$-groups H^*, H with a surjective map such that their derived groups are simply connected (and semisimple), and the abelianization $(H^*)^{\mathrm{ab}}$ is a torus (similarly for H).

Now, $[H^*, H^*] = H^*_{\mathrm{ss}}$ is a simply connected semisimple group and hence it is a product $F_1 \times \cdots \times F_s$ of simply connected $\mathbb{Q}$-simple algebraic groups F_i. Being a factor of $[H^*, H^*] = H^*_{\mathrm{ss}}$, the group $[H, H] = H_{\mathrm{ss}}$ is a product of a (smaller) number of these F_i's. After a renumbering of the indices, we may assume that H_{ss} is a product $F_1 \times \cdots \times F_r$ for some $r \le s$ and the map π on H^*_{ss} is the projection to the first r factors. Hence the image of a congruence subgroup of H^*_{ss} is a congruence subgroup of H_{ss}.

The tori S^* and S have the congruence subgroup property by a result of Chevalley (as already stated at the beginning of this section, this is true for all solvable algebraic groups). Hence the image of a congruence subgroup of S^* is a congruence subgroup of S. We thus need only prove that every subgroup of the reductive group H of the form $\Gamma_1\Gamma_2$, where $\Gamma_1 \subset H_{\mathrm{ss}}$ and $\Gamma_2 \subset S$ are congruence subgroups, is itself a congruence subgroup of H. We use the adelic form of the congruence topology. Suppose K is a compact open subgroup of the $H(\mathbb{A}_f)$ where $\mathbb{A}_f$ is the ring of finite adeles. The image of $H(\mathbb{Q}) \cap K$ under the quotient map $H \to H^{\mathrm{ab}} = S$ is a congruence subgroup in the torus S and hence $H(\mathbb{Q}) \cap K' \subset (H_{\mathrm{ss}}(\mathbb{Q}) \cap K)(S(\mathbb{Q}) \cap K)$ for some possibly smaller open subgroup $K' \subset H(\mathbb{A}_f)$. This proves (iv). □

Note that part (iii) and (iv) prove Proposition 0.1(ii).

2. The arithmetic Chevalley theorem

In this section, we prove Proposition 0.1(i). Assume that $\varphi\colon G_1 \to G_2$ is a surjective morphism of $\mathbb{Q}$-algebraic groups. We are to prove that $\varphi(\Gamma^0)$ contains the commutator subgroup of a congruence subgroup of $G_2(\mathbb{Q})$ containing it.

Before starting on the proof, let us note that in general, the image of a congruence subgroup of $G_1(\mathbb{Z})$ under φ need not be a congruence subgroup of $G_2(\mathbb{Z})$. The following proposition gives a fairly general situation when this happens.

Proposition 2.1. *Let $\pi : G_1 \to G_2$ be a finite covering of semisimple algebraic groups defined over $\mathbb{Q}$ with G_1 simply connected and G_2 not. Assume $G_1(\mathbb{Q})$ is dense in $G_1(\mathbb{A}_f)$. Write K for the kernel of π and K_f for the kernel of the map $G_1(\mathbb{A}_f) \to G_2(\mathbb{A}_f)$. Let Γ be a congruence subgroup of $G_1(\mathbb{Q})$ and H its closure in $G_1(\mathbb{A}_f)$. Then the image $\pi(\Gamma) \subset G_2(\mathbb{Q})$ is a congruence subgroup if and only if $KH \supset K_f$.*

Before proving the proposition, let us note that while K is finite, the group K_f is a product of infinitely many finite abelian groups and that K_f is central in $\overline{G}_1$. This implies:

Corollary 2.2. (i) *There are infinitely many congruence subgroups Γ_i with $\pi(\Gamma_i)$ noncongruence subgroups of unbounded finite index in their congruence closures $\overline{\Gamma}_i$.*

(ii) *For each of these $\Gamma = \Gamma_i$, the image $\pi(\Gamma)$ contains the commutator subgroup $[\overline{\Gamma}, \overline{\Gamma}]$ and is normal in $\overline{\Gamma}$ (with abelian quotient).*

We now prove Proposition 2.1.

Proof. Let G_3 be the image of the rational points of $G_1(\mathbb{Q})$:

$$G_3 = \pi(G_1(\mathbb{Q})) \subset G_2(\mathbb{Q}).$$

Define a subgroup Δ of G_3 to be a *quasicongruence subgroup* if the inverse image $\pi^{-1}(\Delta)$ is a congruence subgroup of $G_1(\mathbb{Q})$. Note that the quasicongruence subgroups of G_3 are exactly the images of congruence subgroups of $G_1(\mathbb{Q})$ by π. It is routine to check that by declaring quasicongruence subgroups to be open, we get the structure of a topological group on G_3. This topology is weaker or equal to the arithmetic topology on G_3. However, it is strictly stronger than the congruence topology on G_3. The last assertion follows from the fact that the completion of $G_3 = G_1(\mathbb{Q})/K(\mathbb{Q})$ is the quotient $\overline{G}_1/K$ where $\overline{G}_1$ is the congruence completion of $G_1(\mathbb{Q})$, whereas the completion of G_3 with respect to the congruence topology is $\overline{G}_1/K_f$.

Now let $\Gamma \subset G_1(\mathbb{Q})$ be a congruence subgroup and $\Delta_1 = \pi(\Gamma)$; let Δ_2 be its congruence closure in G_3. Then both Δ_1 and Δ_2 are open in the quasicongruence topology on G_3. Denote by G_3^* the completion of G_3 with respect to the quasicongruence topology, so $G_3^* = \overline{G}_1/K$ and denote by Δ_1^*, Δ_2^* the closures of Δ_1, Δ_2 in G_3^*. We then have the equalities

$$\Delta_2/\Delta_1 = \Delta_2^*/\Delta_1^*, \quad \Delta_2^* = \Delta_1^* K_f/K.$$

Hence $\Delta_1^* = \Delta_2^*$ if and only if $K\Delta_1^* \supset K_f$. This proves Proposition 2.1.

The proof shows that Δ_1^* is normal in Δ_2^* (since K_f is central) with abelian quotient. The same is true for Δ_1 in Δ_2 and the corollary is also proved. □

To continue with the proof of Proposition 0.1, assume, as we may (by replacing G_1 with the Zariski closure of Γ), that G_1 has no characters defined over $\mathbb{Q}$. For, suppose that G_1 is the Zariski closure of

$\Gamma \subset G_1(\mathbb{Z})$. Let $\chi : G_1 \to \mathbb{G}_m$ be a nontrivial (and therefore surjective) homomorphism defined over $\mathbb{Q}$; then the image of the arithmetic group $G_1(\mathbb{Z})$ in $\mathbb{G}_m(\mathbb{Q})$ is a Zariski dense arithmetic group. However, the only arithmetic groups in $\mathbb{G}_m(\mathbb{Q})$ are finite and cannot be Zariski dense in $\mathbb{G}_m$. Therefore, χ cannot be nontrivial. We can also assume that G_1 is connected.

We start by proving Proposition 0.1 for the case that Γ is a congruence subgroup.

If we write $G_1 = R_1 H_1$ where H_1 is semisimple and R_1 is the radical, we may assume that G_1 is essentially simply connected (Lemma 1.3(iii)), without affecting the hypotheses or the conclusion of Proposition 0.1.

Hence $G_1 = R_1 \rtimes H_1$ is a semidirect product. Then clearly, every congruence subgroup of G_1 contains a congruence subgroup of the form $\Delta \rtimes \Phi$ where $\Delta \subset R_1$ and $\Phi \subset H_1$ are congruence subgroups. Similarly, write $G_2 = R_2 H_2$. Since φ is easily seen to map R_1 onto R_2 and H_1 onto H_2, it is enough to prove the proposition for R_1 and H_1 separately.

We first recall that if G is a solvable linear algebraic group defined over $\mathbb{Q}$ then the congruence subgroup property holds for G, i.e., every arithmetic subgroup of G is a congruence subgroup (for a reference see page 108, last but one paragraph of [Raghunathan 1976] or [Chahal 1980]). Consequently, by Lemma 1.3(ii), the image of a congruence subgroup in R_1 is an arithmetic group in R_2 and hence a congruence subgroup. Thus we dispose of the solvable case.

In the case of semisimple groups, denote by H_2^* by the simply connected cover of H_2. The map $\varphi\colon H_1 \to H_2$ lifts to a map from H_1 to H_2^*. For simply connected semisimple groups, a surjective map from H_1 to H_2^* sends a congruence subgroup to a congruence subgroup by Lemma 1.3(iv).

We are thus reduced to the situation $H_1 = H_2^*$ and $\varphi\colon H_1 \to H_2$ is the simply connected cover of H_2.

By our assumptions, H_1 is now connected, simply connected and semisimple. We claim that for any nontrivial $\mathbb{Q}$-simple factor L of H_1, $L(\mathbb{R})$ is not compact. Otherwise, the image of Γ, the arithmetic group, there is finite and as Γ is Zariski dense, so H_1 is not connected. The strong approximation theorem [Platonov and Rapinchuk 1994, Theorem 7.12] gives now that $H_1(\mathbb{Q})$ is dense in $H_1(\mathbb{A}_f)$. So Proposition 2.1 can be applied to finish the proof of Proposition 0.1 in the case Γ is a congruence subgroup.

We need to show that it is true also for the more general case when Γ is only congruence closed. To this end let us formulate the following proposition which is of independent interest.

Proposition 2.3. *Let* $\Gamma \subseteq \mathrm{GL}_n(\mathbb{Z})$, G *its Zariski closure and* $\mathrm{Der} = [G^0, G^0]$. *Then* Γ *is congruence closed if and only if* $\Gamma \cap \mathrm{Der}$ *is a congruence subgroup of* Der.

Proof. If G^0 has no toral factors, this is proved in [Venkataramana 1999], in fact, in this case a congruence closed Zariski dense subgroup is a congruence subgroup. (Note that this is stated there for general G, but the assumption that there is no toral factor was mistakenly omitted as the proof there shows.)

Now, if there is a toral factor, we can assume G is connected, so $G^{\mathrm{ab}} = V \times S$ where V is unipotent and S a torus. Now $\Gamma \cap [G, G]$ is Zariski dense and congruence closed, so it is a congruence subgroup by [Venkataramana 1999] as before. For the other direction, note that the image of Γ is $U \times S$, being solvable, is always congruence closed, so the proposition follows. □

Now, we can end the proof of Proposition 0.1 for congruence closed subgroups by looking at φ on $G_3 = \overline{\Gamma}$ the Zariski closure of Γ and apply the proof above to $\mathrm{Der}(G_3^0)$. It also proves Proposition 0.2.

Of course, Proposition 2.3 is the general form of the following result from [Venkataramana 1999] (based on [Nori 1987; Weisfeiler 1984]), which is, in fact, the core of Proposition 2.3.

Proposition 2.4. *Suppose $\Gamma \subset G(\mathbb{Z})$ is Zariski dense, G simply connected and Γ a subgroup of $G(\mathbb{Z})$ which is closed in the congruence topology. Then Γ is itself a congruence subgroup.*

3. The Grothendieck closure

The Grothendieck closure of a group Γ.

Definition 3.1. Let $\rho\colon \Gamma \to \mathrm{GL}(V)$ be a representation of Γ on a lattice V in a $\mathbb{Q}$-vector space $V \otimes \mathbb{Q}$. Then we get a continuous homomorphism $\hat{\rho}\colon \hat{\Gamma} \to \mathrm{GL}(\hat{V})$ (where, for a group Δ, $\hat{\Delta}$ denotes its profinite completion) which extends ρ.

Denote by $\mathrm{Cl}_\rho(\Gamma)$ the subgroup of the profinite completion of Γ, which preserves the lattice V: $\mathrm{Cl}_\rho(\Gamma) = \{g \in \hat{\Gamma}\colon \hat{\rho}(g)(V) \subset V\}$. In fact, since $\det(\hat{\rho}(g)) = \pm 1$ for every $g \in \Gamma$ and hence also for every $g \in \hat{\Gamma}$, for $g \in \mathrm{Cl}_g(\Gamma)$, $\hat{\rho}(g)(V) = V$, and hence $\mathrm{Cl}_\rho(\Gamma)$ is a subgroup of $\hat{\Gamma}$. We denote by $\mathrm{Cl}(\Gamma)$ the subgroup

$$\mathrm{Cl}(\Gamma) = \{g \in \hat{\Gamma}\colon \hat{\rho}(g)(V) \subset V \ \forall \text{ lattices } V\}. \tag{3-1}$$

Therefore, $\mathrm{Cl}(\Gamma) = \cap_\rho \mathrm{Cl}_\rho(\Gamma)$ where ρ runs through all integral representations of the group Γ.

Suppose now that V is any finitely generated abelian group (not necessarily a lattice i.e., not necessarily torsion-free) which is also a Γ-module. Then the torsion in V is a (finite) subgroup with finite exponent n say. Then nV is torsion free. Since Γ acts on the finite group V/nV by a finite group via, say, ρ, it follows that $\hat{\Gamma}$ also acts on the finite group V/nV via $\hat{\rho}$. Thus, for $g \in \hat{\Gamma}$ we have $\hat{\rho}(g)(V/nV) = V/nV$. Suppose now that $g \in \mathrm{Cl}(\Gamma)$. Then $g(nV) = nV$ by the definition of $\mathrm{Cl}(\Gamma)$. Hence $g(V)/nV = V/nV$ for $g \in \mathrm{Cl}(\Gamma)$. This is an equality in the quotient group $\hat{V}/nV$. This shows that $g(V) \subset V + nV = V$ which shows that $\mathrm{Cl}(\Gamma)$ preserves *all* finitely generated abelian groups V which are Γ -modules.

By $\mathrm{Cl}_{\mathbb{Z}}(\Gamma)$ we mean the *Grothendieck closure* of the (finitely generated) group Γ. It is essentially a result of [Lubotzky 1980] that the Grothendieck closure $\mathrm{Cl}_{\mathbb{Z}}(\Gamma)$ is the same as the group $\mathrm{Cl}(\Gamma)$ defined above (in [loc. cit.], the group considered was the closure with respect to *all* finitely generated $\mathbb{Z}$ modules which are also Γ modules, whereas we consider only those finitely generated $\mathbb{Z}$ modules which are Γ modules and which are torsion-free; the argument of the preceding paragraph shows that these closures are the same). From now on, we identify the Grothendieck closure $\mathrm{Cl}_{\mathbb{Z}}(\Gamma)$ with the foregoing group $\mathrm{Cl}(\Gamma)$.

Notation. Let Γ be a group, V a finitely generated torsion-free abelian group which is a Γ-module and $\rho\colon \Gamma \to \mathrm{GL}(V)$ the corresponding Γ-action. Denote by G_ρ the Zariski closure of the image $\rho(\Gamma)$ in $\mathrm{GL}(V \otimes \mathbb{Q})$, and G_ρ^0 its connected component of identity. Then both G_ρ, G_ρ^0 are linear algebraic groups defined over $\mathbb{Q}$, and so is $\mathrm{Der}_\rho = [G_\rho^0, G_\rho^0]$.

Let $B = B_\rho(\Gamma)$ denote the subgroup $\hat{\rho}(\hat{\Gamma}) \cap \mathrm{GL}(V)$. Since the profinite topology of $\mathrm{GL}(\hat{V})$ induces the congruence topology on $\mathrm{GL}(V)$, $B_\rho(\Gamma)$ is the congruence closure of $\rho(\Gamma)$ in $\mathrm{GL}(V)$.

We denote by $D = D_\rho(\Gamma)$ the intersection of B with the derived subgroup $\mathrm{Der}_\rho = [G^0, G^0]$. We thus have an exact sequence

$$1 \to D \to B \to A \to 1,$$

where $A = A_\rho(\Gamma)$ is an extension of a finite group G/G^0 by an abelian group (the image of $B \cap G^0$ in the abelianization $(G^0)^{\mathrm{ab}}$ of the connected component G^0).

Simply connected representations.

Definition 3.2. We will say that ρ is *simply connected* if the group $G = G_\rho$ is *essentially simply connected*. That is, if U is the unipotent radical of G, the quotient G^0/U is a product $H \times S$ where H is semisimple and simply connected and S is a torus.

An easy consequence of Lemma 1.2 is that simply connected representations are closed under direct sums.

Lemma 3.3. *Let ρ_1, ρ_2 be two simply connected representations of an abstract group Γ. Then the direct sum $\rho_1 \oplus \rho_2$ is also simply connected.*

We also have:

Lemma 3.4. *Let $\rho\colon \Gamma \to \mathrm{GL}(W)$ be a subrepresentation of a representation $\tau\colon \Gamma \to \mathrm{GL}(V)$ such that both ρ, τ are simply connected. Then the map $r\colon B_\tau(\Gamma) \to B_\rho(\Gamma)$ is surjective.*

Proof. The image of $B_\tau(\Gamma)$ in $B_\rho(\Gamma)$ contains the image of D_τ. By Proposition 2.3, D_τ is a congruence subgroup of the algebraic group Der_τ. The map $\mathrm{Der}_\tau \to \mathrm{Der}_\rho$ is a surjective map between simply connected groups. Therefore, by part (iv) of Lemma 1.3, the image of D_τ is a congruence subgroup F of D_ρ. Now, by Proposition 2.3, $D_\rho \cdot \rho(\Gamma)$ is congruence closed, hence equal to B_ρ which is the congruence closure of $\rho(\Gamma)$ and $B_\tau \to B_\rho$ is surjective. □

Simply connected to general.

Lemma 3.5. *Every (integral) representation $\rho\colon \Gamma \to \mathrm{GL}(W)$ is a subrepresentation of a representation $\tau\colon \Gamma \to \mathrm{GL}(V)$ where τ is simply connected.*

Proof. Let $\rho\colon \Gamma \to \mathrm{GL}(W)$ be a representation. Let Der be the derived subgroup of the identity component of the Zariski closure $H = G_\rho$ of $\rho(\Gamma)$. Then, by Lemma 1.3(iii), there exists a map $H^* \to H^0$ with finite kernel such that H^* is connected and $H^*/U^* = (H^*)_{\mathrm{ss}} \times S^*$ where H^*_{ss} is a simply connected semisimple group. Denote by $W_{\mathbb{Q}}$ the $\mathbb{Q}$-vector space $W \otimes \mathbb{Q}$. By Lemma 1.3(i), $\rho\colon H^0 \to \mathrm{GL}(W_{\mathbb{Q}})$ may be considered as a subrepresentation of a faithful representation $(\theta, E_{\mathbb{Q}})$ of the covering group H^*.

By (ii) of Lemma 1.3, the image of an arithmetic subgroup of H^* is an arithmetic group of H. Moreover, as $H(\mathbb{Z})$ is virtually torsion free, one may choose a normal, torsion-free arithmetic subgroup $\Delta \subset H(\mathbb{Z})$

such that the map $H^* \to H^0$ splits over Δ. In particular, the map $H^* \to H^0$ splits over a normal subgroup N of Γ of finite index. Thus, θ may be considered as a representation of the group N.

Consider the induced representation $\mathrm{Ind}_N^\Gamma(W_\mathbb{Q})$. Since $W_\mathbb{Q}$ is a representation of Γ, it follows that $\mathrm{Ind}_N^\Gamma(W_\mathbb{Q}) = W_\mathbb{Q} \otimes \mathrm{Ind}_N^\Gamma(\mathrm{triv}_N) \supset W_\mathbb{Q}$. Since, by the first paragraph of this proof, $W_\mathbb{Q} \subset E_\mathbb{Q}$ as H^* modules, it follows that $W_\mathbb{Q} \mid_N \subset E_\mathbb{Q}$ and hence $W_\mathbb{Q} \subset \mathrm{Ind}_N^\Gamma(E_\mathbb{Q}) =: V_\mathbb{Q}$. Write $\tau = \mathrm{Ind}_N^\Gamma(E_\mathbb{Q})$ for the representation of Γ on $V_\mathbb{Q}$. The normality of N in Γ implies that the restriction representation $\tau \mid_N$ is contained in a direct sum of the N-representations $n \to \theta(\gamma n \gamma^{-1})$ as γ varies over the finite set Γ/N.

Write $G_{\theta|_N}$ for the Zariski closure of the image $\theta(N)$. Since $G_{\theta|_N}$ has H^* as its Zariski closure and the group H^*_{ss} is simply connected, each θ composed with conjugation by γ is a simply connected representation of N. It follows from Lemma 3.3 that $\tau \mid_N$ is simply connected. Since simple connectedness of a representation is the same for subgroups of finite index, it follows that τ, as a representation of Γ, is simply connected.

We have now proved that there exists Γ-equivariant embedding of the module $(\rho, W_\mathbb{Q})$ into $(\tau, V_\mathbb{Q})$ where W, V are lattices in the $\mathbb{Q}$-vector spaces $W_\mathbb{Q}, V_\mathbb{Q}$. A basis of the lattice W is then a $\mathbb{Q}$-linear combination of a basis of V; the finite generation of W then implies that there exists an integer m such that $mW \subset V$, and this inclusion is an embedding of Γ-modules. Clearly, the module (ρ, W) is isomorphic to (ρ, mW) the isomorphism given by multiplication by m. Hence the lemma follows. □

The following is the main technical result of this section, from which the main results of this paper are derived:

Proposition 3.6. *The group* $\mathrm{Cl}(\Gamma)$ *is the inverse limit of the groups* $B_\rho(\Gamma)$ *where* ρ *runs through simply connected representations and* $B_\rho(\Gamma)$ *is the congruence closure of* $\rho(\Gamma)$. *Moreover, if* $\rho\colon \Gamma \to \mathrm{GL}(W)$ *is simply connected, then the map* $\mathrm{Cl}(\Gamma) \to B_\rho(\Gamma)$ *is surjective.*

Proof. Denote temporarily by $\mathrm{Cl}(\Gamma)_{sc}$ the subgroup of elements of $\hat{\Gamma}$ which stabilize the lattice V for all *simply connected* representations (τ, V). Let W be an arbitrary finitely generated torsion-free lattice which is also a Γ-module; denote by ρ the action of Γ on W.

By Lemma 3.5, there exists a simply connected representation (τ, V) which contains (ρ, W). If $g \in \mathrm{Cl}(\Gamma)_{sc}$, then $\hat{\tau}(g)(V) \subset V$; since Γ is dense in $\hat{G}$ and stabilizes W, it follows that for all $x \in \hat{\Gamma}$, $\hat{\tau}(x)(\hat{W}) \subset \hat{W}$; in particular, for $g \in \mathrm{Cl}(\Gamma)_{sc}$, $\hat{\rho}(g)(W) = \hat{\tau}(g)(W) \subset \hat{W} \cap V = W$. Thus, $\mathrm{Cl}(\Gamma)_{sc} \subset \mathrm{Cl}(\Gamma)$.

The group $\mathrm{Cl}(\Gamma)$ is, by definition, the set of all elements g of the profinite completion $\hat{\Gamma}$ which stabilize all Γ stable torsion free lattices. It follows in particular, that these elements g stabilize all Γ-stable lattices V associated to simply connected representations (τ, V); hence $\mathrm{Cl}(\Gamma) \subset \mathrm{Cl}(\Gamma)_{sc}$. The preceding paragraph now implies that $\mathrm{Cl}(\Gamma) = \mathrm{Cl}(\Gamma)_{sc}$. This proves the first part of the proposition (see (0-2)).

We can enumerate all the simply connected integral representations ρ, since Γ is finitely generated. Write $\rho_1, \rho_2, \ldots, \rho_n \ldots,$ for the sequence of simply connected representations of Γ. Write τ_n for the direct sum $\rho_1 \oplus \rho_2 \oplus \cdots \oplus \rho_n$. Then $\tau_n \subset \tau_{n+1}$ and by Lemma 3.3 each τ is simply connected; moreover, the simply connected representation ρ_n is contained in τ_n.

By Lemma 3.4, it follows that $\mathrm{Cl}(\Gamma)$ is the inverse limit of the *totally ordered family* $B_{\tau_n}(\Gamma)$; moreover, $B_{\tau_{n+1}}(\Gamma)$ maps *onto* $B_{\tau_n}(\Gamma)$. By taking inverse limits, it follows that $\mathrm{Cl}(\Gamma)$ maps *onto* the group $B_{\tau_n}(\Gamma)$ for every n. It follows, again from Lemma 3.4, that every $B_{\rho_n}(\Gamma)$ is a homomorphic image of $B_{\tau_n}(\Gamma)$ and hence of $\mathrm{Cl}(\Gamma)$. This proves the second part of the proposition. □

Definition 3.7. Let Γ be a finitely generated group. We say that Γ is FAb if the abelianization Δ^{ab} is finite for every finite index subgroup $\Delta \subset \Gamma$.

Corollary 3.8. *If Γ is* FAb *then for every simply connected representation ρ, the congruence closure $B_\rho(\Gamma)$ of $\rho(\Gamma)$ is a congruence subgroup and* $\mathrm{Cl}(\Gamma)$ *is an inverse limit over a totally ordered set τ_n of simply connected representations of Γ, of congruence groups B_n in groups $G_n = G_{\tau_n}$ with G_n^0 simply connected. Moreover, the maps $B_{n+1} \to B_n$ are surjective. Hence the maps* $\mathrm{Cl}(\Gamma) \to B_n$ *are all surjective.*

Proof. If $\rho\colon \Gamma \to \mathrm{GL}(V)$ is a simply connected representation, then for a finite index subgroup Γ^0 the image $\rho(\Gamma^0)$ has connected Zariski closure, and by assumption, $G^0/U = H \times S$ where S is a torus and H is simply connected semisimple. Since the group Γ is FAb it follows that $S = 1$ and hence $G^0 = \mathrm{Der}(G^0)$. Now Proposition 2.4 implies that $B_\rho(\Gamma)$ is a congruence subgroup of $G_\rho(V)$. The Corollary is now immediate from the Proposition 3.6. We take $B_n = B_{\tau_n}$ in the proof of the proposition. □

We can now prove Theorem 0.5. Let us first prove the direction claiming that the congruence subgroup property implies $\mathrm{Cl}(\Gamma) = \Gamma$. This was proved for arithmetic groups Γ by Grothendieck, and we follow here the proof in [Lubotzky 1980] which works for general Γ. Indeed, if $\rho\colon \Gamma \to \mathrm{GL}_n(\mathbb{Z})$ is a faithful simply connected representation such that $\rho(\Gamma)$ satisfies the congruence subgroup property, then it means that the map $\hat{\rho}\colon \hat{\Gamma} \to \mathrm{GL}_n(\hat{\mathbb{Z}})$ is injective. Now $\rho(\mathrm{Cl}(\Gamma)) \subseteq \mathrm{GL}_n(\mathbb{Z}) \cap \hat{\rho}(\hat{\Gamma})$, but the last is exactly the congruence closure of $\rho(\Gamma)$. By our assumption, $\rho(\Gamma)$ is congruence closed, so it is equal to $\rho(\Gamma)$. So in summary $\hat{\rho}(\Gamma) \subset \hat{\rho}(\mathrm{Cl}(\Gamma)) \subseteq \rho(\Gamma) = \hat{\rho}(\Gamma)$. As $\hat{\rho}$ is injective, $\Gamma = \mathrm{Cl}(\Gamma)$.

In the opposite direction, assume $\mathrm{Cl}(\Gamma) = \Gamma$. By the description of $\mathrm{Cl}(\Gamma)$ in (0.1) or in (3.1), it follows that for every finite index subgroup Γ' of Γ, $\mathrm{Cl}(\Gamma') = \Gamma'$ (see [Lubotzky 1980, Proposition 4.4]). Now, if ρ is a faithful simply connected representation of Γ, it is also such for Γ' and by Proposition 3.6, $\rho(\mathrm{Cl}(\Gamma'))$ is congruence closed. In our case it means that for every finite index subgroup Γ', $\rho(\Gamma')$ is congruence closed, i.e., $\rho(\Gamma)$ has the congruence subgroup property.

4. Thin groups

Let Γ be a finitely generated $\mathbb{Z}$-linear group, i.e., $\Gamma \subset \mathrm{GL}_n(\mathbb{Z})$, for some n. Let G be its Zariski closure in $\mathrm{GL}_n(\mathbb{C})$ and $\Delta = G \cap \mathrm{GL}_n(\mathbb{Z})$. We say that Γ is a *thin* subgroup of G if $[\Delta\colon \Gamma] = \infty$, otherwise Γ is an arithmetic subgroup of G. In general, given Γ, (say, given by a set of generators) it is a difficult question to determine if Γ is thin or arithmetic. Our next result gives, still, a group theoretic characterization for the *abstract* group Γ to be thin. But first a warning: an abstract group can sometimes appear as an arithmetic subgroup and sometimes as a thin subgroup. For example, the free group on two generators $F = F_2$ is a finite index subgroup of $\mathrm{SL}_2(\mathbb{Z})$, and so, arithmetic. But at the same time, by a well-known result of Tits [1972] asserting that $\mathrm{SL}_n(\mathbb{Z})$ contains a copy of F which is Zariski dense in SL_n; it is also thin. To be precise, let us define:

Definition 4.1. A finitely generated $\mathbb{Z}$-linear group Γ is called a *thin group* if it has a faithful representation $\rho\colon \Gamma \to \mathrm{GL}_n(\mathbb{Z})$ for some $n \in \mathbb{Z}$, such that $\rho(\Gamma)$ is of infinite index in $\overline{\rho(\Gamma)}^Z \cap \mathrm{GL}_n(\mathbb{Z})$ where $\overline{\rho(\Gamma)}^Z$ is the Zariski closure of Γ in GL_n. Such a ρ will be called a thin representation of Γ.

We have assumed that $i\colon \Gamma \subset \mathrm{GL}_n(\mathbb{Z})$. Assume also, as we may (see Lemma 3.5) that the representation i is simply connected. By Proposition 3.6, the group $\mathrm{Cl}(\Gamma)$ is the subgroup of $\hat{\Gamma}$ which preserves the lattices V_n for a totally ordered set (with respect to the relation of being a subrepresentation) of faithful simply connected integral representations (ρ_n, V_n) of Γ with the maps $\mathrm{Cl}(\Gamma) \to B_n$ being surjective, where B_n is the congruence closure of $\rho_n(\Gamma)$ in $\mathrm{GL}(V_n)$. Hence, $\mathrm{Cl}(\Gamma)$ is the inverse limit (as n varies) of the congruence closed subgroups B_n and Γ is the inverse limit of the images $\rho_n(\Gamma)$. Equip $B_n/\rho_n(\Gamma)$ with the discrete topology. Consequently, $\mathrm{Cl}(\Gamma)/\Gamma$ is a closed subspace of the Tychonov product $\prod_n(B_n/\rho_n(\Gamma))$. This is the topology on $\mathrm{Cl}(\Gamma)/\Gamma$ considered in the following theorem.

Theorem 4.2. *Let Γ be a finitely generated $\mathbb{Z}$-linear group, i.e., $\Gamma \subset \mathrm{GL}_m(\mathbb{Z})$ for some n. Then Γ is* not *a thin group if and only if Γ satisfies both of the following two properties*:

(a) *Γ is a* FAb *group* (*i.e., for every finite index subgroup Λ of Γ, $\Lambda/[\Lambda, \Lambda]$ is finite*).

(b) *The group* $\mathrm{Cl}(\Gamma)/\Gamma$ *is compact.*

Proof. Assume first that Γ is a thin group. If Γ is not FAb we are done. So, assume Γ is FAb. We must now prove that $\mathrm{Cl}(\Gamma)/\Gamma$ is not compact. We know that Γ has a faithful thin representation $\rho\colon \Gamma \to \mathrm{GL}_n(\mathbb{Z})$ which in addition, is simply connected. This induces a surjective map (see Proposition 3.6) $\mathrm{Cl}(\Gamma) \to B_\rho(\Gamma)$ where $B_\rho(\Gamma)$ is the congruence closure of $\rho(\Gamma)$ in $\mathrm{GL}_n(\mathbb{Z})$. As Γ is FAb, $B_\rho(\Gamma)$ is a congruence subgroup, by Corollary 3.8. But as ρ is thin, $\rho(\Gamma)$ has infinite index in $B_\rho(\Gamma)$. Thus, $\mathrm{Cl}(\Gamma)/\Gamma$ is mapped *onto* the discrete infinite quotient space $B_\rho(\Gamma)/\rho(\Gamma)$. Hence $\mathrm{Cl}(\Gamma)/\Gamma$ is not compact.

Assume now Γ is not a thin group. This implies that for every faithful integral representation $\rho(\Gamma)$ is of finite index in its integral Zariski closure. We claim that $\Gamma/[\Gamma, \Gamma]$ is finite. Otherwise, as Γ is finitely generated, Γ is mapped on $\mathbb{Z}$. The group $\mathbb{Z}$ has a Zariski dense integral representation τ into $\mathbb{G}_a \times S$ where S is a torus; take any integral matrix $g \in \mathrm{SL}_n(\mathbb{Z})$ which is neither semisimple nor unipotent, whose semisimple part has infinite order. Then both the unipotent and semisimple part of the Zariski closure H of $\tau(\mathbb{Z})$ are nontrivial and $H(\mathbb{Z})$ cannot contain $\tau(\mathbb{Z})$ as a subgroup of finite index since $H(\mathbb{Z})$ is commensurable to $\mathbb{G}_a(\mathbb{Z}) \times S(\mathbb{Z})$ and both factors are nontrivial and infinite. The representation $\rho \times \tau$ (where ρ is any faithful integral representation of Γ) will give a thin representation of Γ. This proves that $\Gamma/[\Gamma, \Gamma]$ is finite. A similar argument (using an induced representation) works for every finite index subgroup, hence Γ satisfies FAb.

We now prove that $\mathrm{Cl}(\Gamma)/\Gamma$ is compact. We already know that Γ is FAb, so by Corollary 3.8, $\mathrm{Cl}(\Gamma) = \varprojlim B_{\rho_n}(\Gamma)$ when $B_n = B_{\rho_n}(\Gamma)$ are congruence groups with surjective homomorphisms $B_{n+1} \to B_n$. Note that as Γ has a faithful integral representation, we can assume that all the representations ρ_n in the sequence are faithful and

$$\Gamma = \varprojlim_{n} \rho_n(\Gamma). \tag{4-1}$$

This implies that $\mathrm{Cl}(\Gamma)/\Gamma = \varprojlim_n B_n/\rho_n(\Gamma)$. Now, by our assumption, each $\rho_n(\Gamma)$ is of finite index in $B_n = B_{\rho_n}(\Gamma)$. So $\mathrm{Cl}(\Gamma)/\Gamma$ is an inverse limit of finite sets and hence compact. □

5. Grothendieck closure and super-rigidity

Let Γ be a finitely generated group. We say that Γ is *integral super-rigid* if there exists an algebraic group $G \subseteq \mathrm{GL}_m(\mathbb{C})$ and an embedding $i\colon \Gamma_0 \mapsto G$ of a finite index subgroup Γ_0 of Γ, such that for every integral representation $\rho\colon \Gamma \to \mathrm{GL}_n(\mathbb{Z})$, there exists an algebraic representation $\tilde{\rho}\colon G \to \mathrm{GL}_n(\mathbb{C})$ such that ρ and $\tilde{\rho}$ agree on some finite index subgroup of Γ_0. Note: Γ is integral super-rigid if and only if a finite index subgroup of Γ is integral super-rigid.

Examples of such super-rigid groups are, first of all, the irreducible (arithmetic) lattices in high rank semisimple Lie groups, but also the (arithmetic) lattices in the rank one simple Lie groups $\mathrm{Sp}(n, 1)$ and $\mathbb{F}_4^{-20}$ (see [Margulis 1991; Corlette 1992; Gromov and Schoen 1992]). But [Bass and Lubotzky 2000] shows that there are such groups which are thin groups.

Now, let Γ be a subgroup of $\mathrm{GL}_m(\mathbb{Z})$, whose Zariski closure is essentially simply connected. We say that Γ satisfies the *congruence subgroup property* (CSP) if the natural extension of $i\colon \Gamma \to \mathrm{GL}_m(\mathbb{Z})$ to $\hat{\Gamma}$, i.e., $\tilde{i}\colon \hat{\Gamma} \to \mathrm{GL}_m(\hat{\mathbb{Z}})$ has finite kernel.

Theorem 5.1. *Let* $\Gamma \subseteq \mathrm{GL}_m(\mathbb{Z})$ *be a finitely generated subgroup satisfying* (FAb). *Then*:

(a) $\mathrm{Cl}(\Gamma)/\Gamma$ *is compact if and only if* Γ *is an arithmetic group which is integral super-rigid.*

(b) $\mathrm{Cl}(\Gamma)/\Gamma$ *is finite if and only if* Γ *is an arithmetic group satisfying the congruence subgroup property.*

Remarks. (a) The finiteness of $\mathrm{Cl}(\Gamma)/\Gamma$ implies, in particular, its compactness, so Theorem 5.1 recovers the well-known fact (see [Bass et al. 1967; Raghunathan 1976]) that the congruence subgroup property implies super-rigidity.

(b) As explained in Section 2 (based on [Serre 1968]) the simple connectedness is a necessary condition for the CSP to hold. But by Lemma 3.5, if Γ has any embedding into $\mathrm{GL}_n(\mathbb{Z})$ for some n, it also has a simply connected one.

We now prove Theorem 5.1.

Proof. Assume first $\mathrm{Cl}(\Gamma)/\Gamma$ is compact in which case, by Theorem 4.2, Γ must be an arithmetic subgroup of some algebraic group G. Without loss of generality (using Lemma 3.5) we can assume that G is connected and simply connected, call this representation $\rho\colon \Gamma \to G$. Let θ be any other representation of Γ.

Let $\tau = \rho \oplus \theta$ be the direct sum. The group G_τ is a subgroup of $G_\rho \times G_\theta$ with surjective projections. Since both τ and ρ are embeddings of the group Γ, and Γ does not have thin representations, it follows (from Corollary 3.8) that the projection $\pi : G_\tau \to G_\rho$ yields an isomorphism of the arithmetic groups $\tau(\Gamma) \subset G_\tau(\mathbb{Z})$ and $\rho(\Gamma) \subset G_\rho(\mathbb{Z})$.

Assume, as we may, that Γ is torsion-free and Γ is an arithmetic group. Every arithmetic group in $G_\tau(\mathbb{Z})$ is virtually a product of the form $U_\tau(\mathbb{Z}) \rtimes H_\tau(\mathbb{Z})$ where U_τ and H_τ are the unipotent and semisimple parts of G_τ respectively (note that G_τ^0 cannot have torus as quotient since Γ is FAb). Hence $\Gamma \cap U_\tau(\mathbb{Z})$ may also be described as the virtually maximal normal nilpotent subgroup of Γ. Similarly for $\Gamma \cap U_\rho(\mathbb{Z})$. This proves that the groups U_τ and U_ρ have isomorphic arithmetic groups which proves that $\pi : U_\tau \to U_\rho$ is an isomorphism. Otherwise $\mathrm{Ker}(\pi)$, which is a $\mathbb{Q}$-defined normal subgroup of U_τ, would have an infinite intersection with the arithmetic group $\Gamma \cap U_\tau$.

Therefore, the arithmetic groups in H_τ and H_ρ are isomorphic and the isomorphism is induced by the projection $H_\tau \to H_\rho$. Since H_ρ is simply connected by assumption, and is a factor of H_τ, it follows that H_τ is a product $H_\rho H$ where H is a semisimple group defined over $\mathbb{Q}$ with $H(\mathbb{Z})$ Zariski dense in H. But the isomorphism of the arithmetic groups in H_τ and H_ρ then shows that the group $H(\mathbb{Z})$ is finite which means that H is finite. Therefore, $\pi : H_\tau^0 \to H_\rho$ is an isomorphism and so the map $G_\tau^0 \to G_\rho$ is also an isomorphism since it is a surjective morphism between groups of the same dimension, and since G_ρ is simply connected.

This proves that Γ is a super-rigid group.

In [Lubotzky 1980], it was proved that if Γ satisfies super-rigidity in some simply connected group G, then (up to finite index) $\mathrm{Cl}(\Gamma)/\Gamma$ is in one-to-one correspondence with

$$C(\Gamma) = \mathrm{Ker}(\hat{\Gamma} \to G(\hat{\mathbb{Z}})).$$

This finishes the proof of both parts (a) and (b). □

Remark. In the situation of Theorem 5.1, Γ is an arithmetic group satisfying super-rigidity. The difference between parts (a) and (b), is whether Γ also satisfies CSP. As of now, there is no known arithmetic group (in a simply connected group) which satisfies super-rigidity without satisfying CSP. The conjecture of Serre about the congruence subgroup problem predicts that arithmetic lattices in rank one Lie groups fail to have CSP. These include Lie groups like $\mathrm{Sp}(n, 1)$ and $\mathbb{F}_4^{(-20)}$ for which super-rigidity was shown (after Serre had made his conjecture). Potentially, the arithmetic subgroups of these groups can have $\mathrm{Cl}(\Gamma)/\Gamma$ compact and not finite. But (some) experts seem to believe now that these groups do satisfy CSP. Anyway as of now, we do not know any subgroup Γ of $\mathrm{GL}_n(\mathbb{Z})$ with $\mathrm{Cl}(\Gamma)/\Gamma$ compact and not finite.

Acknowledgments

T.N. Venkataramana thanks the Math Department of the Hebrew University for great hospitality while a major part of this work was done. He would also like to thank the JC Bose fellowship (SR/S2/JCB-22/2017) for support during the period 2013–2018.

The authors thank the Math Department of the University of Marseilles, and the conference at Oberwolfach, where the work was completed. We would especially like to thank Bertrand Remy for many interesting discussions and for his warm hospitality.

A. Lubotzky is indebted for support from the NSF, the ISF and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No. 692854).

References

[Bass and Lubotzky 2000] H. Bass and A. Lubotzky, "Nonarithmetic superrigid groups: counterexamples to Platonov's conjecture", *Ann. of Math.* (2) **151**:3 (2000), 1151–1173. MR Zbl

[Bass et al. 1967] H. Bass, J. Milnor, and J.-P. Serre, "Solution of the congruence subgroup problem for SL_n ($n \geq 3$) and Sp_{2n} ($n \geq 2$)", *Inst. Hautes Études Sci. Publ. Math.* **33** (1967), 59–137. MR Zbl

[Bridson and Grunewald 2004] M. R. Bridson and F. J. Grunewald, "Grothendieck's problems concerning profinite completions and representations of groups", *Ann. of Math.* (2) **160**:1 (2004), 359–373. MR Zbl

[Chahal 1980] J. S. Chahal, "Solution of the congruence subgroup problem for solvable algebraic groups", *Nagoya Math. J.* **79** (1980), 141–144. MR Zbl

[Corlette 1992] K. Corlette, "Archimedean superrigidity and hyperbolic geometry", *Ann. of Math.* (2) **135**:1 (1992), 165–182. MR Zbl

[Gromov and Schoen 1992] M. Gromov and R. Schoen, "Harmonic maps into singular spaces and p-adic superrigidity for lattices in groups of rank one", *Inst. Hautes Études Sci. Publ. Math.* **76** (1992), 165–246. MR Zbl

[Grothendieck 1970] A. Grothendieck, "Représentations linéaires et compactification profinie des groupes discrets", *Manuscripta Math.* **2** (1970), 375–396. MR Zbl

[Kontorovich et al. 2019] A. Kontorovich, D. D. Long, A. Lubotzky, and A. W. Reid, "What is... a thin group?", *Notices Amer. Math. Soc.* **66**:6 (2019), 905–910.

[Lubotzky 1980] A. Lubotzky, "Tannaka duality for discrete groups", *Amer. J. Math.* **102**:4 (1980), 663–689. MR Zbl

[Margulis 1991] G. A. Margulis, *Discrete subgroups of semisimple Lie groups*, Ergebnisse der Mathematik (3) **17**, Springer, 1991. MR Zbl

[Nori 1987] M. V. Nori, "On subgroups of $\mathrm{GL}_n(\mathbb{F}_p)$", *Invent. Math.* **88**:2 (1987), 257–275. MR Zbl

[Platonov and Rapinchuk 1994] V. Platonov and A. Rapinchuk, *Algebraic groups and number theory*, Pure and Appl. Math. **139**, Academic Press, Boston, 1994. MR Zbl

[Platonov and Tavgen 1986] V. P. Platonov and O. I. Tavgen, "On the Grothendieck problem of profinite completions of groups", *Dokl. Akad. Nauk SSSR* **288**:5 (1986), 1054–1058. In Russian; translated in *Soviet Math. Dokl.* **33**:3 (1986), 822–825. MR Zbl

[Pyber 2004] L. Pyber, "Groups of intermediate subgroup growth and a problem of Grothendieck", *Duke Math. J.* **121**:1 (2004), 169–188. MR Zbl

[Raghunathan 1976] M. S. Raghunathan, "On the congruence subgroup problem", *Inst. Hautes Études Sci. Publ. Math.* **46** (1976), 107–161. MR Zbl

[Raghunathan 1991] M. S. Raghunathan, "The congruence subgroup problem", pp. 465–494 in *Proceedings of the Hyderabad Conference on Algebraic Groups* (Hyderabad, 1989), edited by S. Ramanan et al., Manoj Prakashan, Madras, 1991. MR Zbl

[Sarnak 2014] P. Sarnak, "Notes on thin matrix groups", pp. 343–362 in *Thin groups and superstrong approximation* (Berkeley, 2012), edited by E. Breuillard and H. Oh, Math. Sci. Res. Inst. Publ. **61**, Cambridge Univ. Press, 2014. MR Zbl

[Serre 1968] J.-P. Serre, "Groupes de congruence (d'après H. Bass, H. Matsumoto, J. Mennicke, J. Milnor, C. Moore)", exposé 330 in *Séminaire Bourbaki,* 1966/1967, W. A. Benjamin, Amsterdam, 1968. Reprinted as pp. 275–291 in *Séminaire Bourbaki* **10**, Soc. Math. France, Paris, 1995. MR Zbl

[Tits 1972] J. Tits, "Free subgroups in linear groups", *J. Algebra* **20** (1972), 250–270. MR Zbl

[Venkataramana 1999] T. N. Venkataramana, "A remark on extended congruence subgroups", *Int. Math. Res. Not.* **1999**:15 (1999), 835–838. MR Zbl

[Weisfeiler 1984] B. Weisfeiler, "Strong approximation for Zariski-dense subgroups of semi-simple algebraic groups", *Ann. of Math.* (2) **120**:2 (1984), 271–315. MR Zbl

Communicated by Peter Sarnak
Received 2018-05-08 Revised 2019-01-20 Accepted 2019-03-08

alex.lubotzky@mail.huji.ac.il *Einstein Institute of Mathematics, The Hebrew University of Jerusalem, Israel*

venky@math.tifr.res.in *School of Mathematics, Tata Institute of Fundamental Research, Mumbai, India*

mathematical sciences publishers msp

ALGEBRA AND NUMBER THEORY 13:6 (2019)
dx.doi.org/10.2140/ant.2019.13.1299

On the ramified class field theory of relative curves

Quentin Guignard

We generalize Deligne's approach to tame geometric class field theory to the case of a relative curve, with arbitrary ramification.

1. Introduction

Let $X \to S$ be a relative curve, i.e., a smooth morphism of schemes of relative dimension 1, with connected geometric fibers, which is Zariski-locally projective over S. Let $Y \hookrightarrow X$ be a relative effective Cartier divisor over S (see Section 4.10), and let U be the complement of Y in X.

The pairs $(\mathcal{L}, \alpha)$, where $\mathcal{L}$ is an invertible $\mathcal{O}_X$-module and α is a rigidification of $\mathcal{L}$ along Y, are parametrized by an S-group scheme $\mathrm{Pic}_S(X, Y)$, the relative rigidified Picard scheme (see Proposition 4.8). The Abel–Jacobi morphism

$$\Phi : U \to \mathrm{Pic}_S(X, Y)$$

is the morphism which sends a section x of U to the pair $(\mathcal{O}(x), 1)$, see Proposition 4.14. We prove the following relative version of the main theorem of geometric global class field theory:

Theorem 1.1 (Theorem 5.3). *Let Λ be a finite ring of cardinality invertible on S, and let $\mathcal{F}$ be an étale sheaf of Λ-modules, locally free of rank 1 on U, with ramification bounded by Y* (see Definition 5.2). *There exists a unique* (*up to isomorphism*) *multiplicative étale sheaf of Λ-modules $\mathcal{G}$ on* $\mathrm{Pic}_S(X, Y)$, *locally free of rank* 1, *such that the pullback of $\mathcal{G}$ by Φ is isomorphic to $\mathcal{F}$.*

MSC2010: 11G45.
Keywords: geometric class field theory, global class field theory, ramification.

The notion of multiplicative locally free Λ-module of rank 1 is defined in Definition 2.5, and it corresponds to isogenies $G \to \mathrm{Pic}_S(X, Y)$ with constant kernel $\Lambda^\times$. We restrict ourselves in this article to $\Lambda^\times$-torsors, with Λ as in Theorem 1.1, in order to simplify the exposition, since we are able to apply directly our main descent tool in this context, namely Lemma 5.9. However, the latter lemma, and hence Theorem 1.1 can be extended to G-torsors, where G is an arbitrary locally constant finite abelian group on $S_{\text{ét}}$.

The case where S is the spectrum of a perfect field is originally due to Serre [1959] and Lang [1956, §6]. Their proof relies on the Albanese property of Rosenlicht's generalized Jacobians [Rosenlicht 1954]. A similar proof was sketched in a letter of 1974 from Deligne to Serre [Deligne 2001]. However, a more geometric proof was given by Deligne in the tamely ramified case; an account of his proof in the unramified case over a finite field can be found in [Laumon 1990, Section 2]. We generalize the latter approach by Deligne to allow arbitrary ramification and an arbitrary base S. This generalization is inspired by notes by Alain Genestier (unpublished) on arithmetic global class field theory.

Deligne's approach has the advantage over Serre and Lang's to yield an explicit geometric construction of the isogeny over $\mathrm{Pic}_S(X, Y)$ corresponding to a local system of rank 1 over U. This feature of Deligne's approach carries over to ours, and is in fact crucial in order to handle the case of an arbitrary base S.

The author learned during the preparation of this manuscript that Daichi Takeuchi had independently obtained a different proof of Theorem 1.1 in the case where S is the spectrum of a perfect field, also by generalizing Deligne's approach to handle arbitrary ramification. See [Takeuchi 2019].

Notation and conventions. We fix a universe $\mathcal{U}$ [SGA 4_3 1973, I.0]. Throughout this paper, all sets are assumed to belong to $\mathcal{U}$ and we will use the term "topos" as a shorthand for "$\mathcal{U}$-topos" [SGA 4_3 1973, IV.1.1]. The category of sets belonging to $\mathcal{U}$ is simply denoted by Sets.

For any integers e, d we denote by $[\![e, d]\!]$ the set of integers n such that $e \leq n \leq d$ and by $\mathfrak{S}_d$ the group of bijections of $[\![1, d]\!]$ onto itself.

In this paper, all rings are unital and commutative. For any ring A, we denote by Alg_A the category of A-algebras. For any scheme S, we denote by $\mathrm{Sch}_{/S}$ the category of S-schemes. We denote by $S_{\text{ét}}$ (resp. $S_{\text{Ét}}$) the small étale topos (resp. big étale topos) of a scheme S, i.e., the topos of sheaves of sets for the étale topology [SGA 4_3 1973, VII.1.2] on the category of étale S-schemes (resp. on $\mathrm{Sch}_{/S}$), and by S_{Fppf} the big fppf topos of S, i.e., the topos of sheaves of sets for the fppf topology on $\mathrm{Sch}_{/S}$ [SGA 4_3 1973, VII.4.2]. If $f : X \to S$ is a morphism of schemes, then we denote by (f^{-1}, f_*) the induced morphism of topos from $X_{\text{Ét}}$ to $S_{\text{Ét}}$. The symbol f^* will exclusively denote the pullback functor from $\mathcal{O}_S$-modules to $\mathcal{O}_X$-modules.

2. Preliminaries

2.1. Let E be a topos and let G be an abelian group in E. We denote by GE the category of objects of E endowed with a left action of G. If X is an object of E, we denote by $E_{/X}$ the topos of X-objects in E. If X is considered as an object of GE by endowing it with the trivial left G-action, then we have $(GE)_{/X} = G(E_{/X})$ and this category will be simply denoted by $GE_{/X}$.

Definition 2.2. A *G-torsor* over an object X of E is an object P of $GE_{/X}$ such that $P \to X$ is an epimorphism and the morphism

$$G \times P \to P \times_X P, \quad (g, p) \mapsto (g \cdot p, p)$$

is an isomorphism. We denote by $\mathrm{Tors}(X, G)$ the full subcategory of $GE_{/X}$ whose objects are the G-torsors over X. If $f : Y \to X$ is a morphism in E, we denote by $f^{-1} : \mathrm{Tors}(X, G) \to \mathrm{Tors}(Y, G)$ the functor which associates $f^{-1}P = P \times_{X,f} Y$ to a G-torsor P over X.

The category $\mathrm{Tors}(X, G)$ is monoidal, with product

$$P_1 \otimes P_2 = G_2 \setminus P_1 \times_X P_2,$$

where G_2 is the kernel of the multiplication morphism $G \times G \to G$, and where $G_2 \hookrightarrow G \times G$ acts diagonally on $P_1 \times_X P_2$. The neutral element for this product is the trivial G-torsor over X, namely $G \times X$, and each G-torsor P over X is invertible with respect to $\otimes$, with inverse given by

$$P^{-1} = \underline{\mathrm{Hom}}_{GE_{/X}}(P, G \times X),$$

where $\underline{\mathrm{Hom}}_{GE_{/X}}$ denotes the internal Hom functor in $GE_{/X}$.

Example 2.3. If $G = \Lambda^\times$ for some ring Λ in E, then the monoidal category $\mathrm{Tors}(X, G)$ is equivalent to the groupoid of locally free Λ-modules of rank 1 in $E_{/X}$. The equivalence is given by the functor which sends an object P of $\mathrm{Tors}(X, G)$ to the Λ-module $G \setminus (\Lambda \times P)$, where the action of $G = \Lambda^\times$ on $\Lambda \times P$ is given by the formula $g \cdot (\lambda, p) = (g\lambda, g \cdot p)$. The functor which sends a locally free Λ-module M of rank 1 of $E_{/X}$ to the G-torsor of isomorphisms of Λ-modules from M to Λ defines a quasiinverse to the latter functor.

2.4. Let E be a topos, and let us denote by 1 its terminal object. Let us consider an exact sequence

$$1 \to G \xrightarrow{i} P \xrightarrow{r} Q \to 1$$

of abelian groups in E. The morphism

$$G \times P \to P \times_Q P, \quad (g, p) \mapsto (i(g) + p, p)$$

is an isomorphism, so that P is a G-torsor over Q. Moreover, the multiplication morphism

$$P \times P \to P$$

factors though $G_2 \setminus P \times P$, where $G_2 \hookrightarrow G \times G$ is the kernel of the multiplication morphism of G, acting diagonally on $P \times P$. We thus obtain a morphism

$$p_1^{-1} P \otimes p_2^{-1} P \to m^{-1} P$$

of G-torsors over $Q \times Q$, where p_1 and p_2 are the canonical projections and m is the multiplication morphism of Q.

The following definition is inspired by [Moret-Bailly 1985, I.2.3]:

Definition 2.5. Let G be an abelian group of E and let Q be a commutative semigroup of E (with or without identity). Let $m : Q \times Q \to Q$ be the multiplication morphism of Q. A *multiplicative G-torsor* over Q is a G-torsor $P \to Q$, together with an isomorphism

$$\theta : p_1^{-1}P \otimes p_2^{-1}P \to m^{-1}P$$

of G-torsors over $Q \times Q$, where p_1 and p_2 are the canonical projections, which satisfy the following two properties.

▹ *Symmetry*: If σ is the involution of $Q \times Q$ which switches the two factors, then the isomorphism

$$p_2^{-1}P \otimes p_1^{-1}P \to \sigma^{-1}(p_1^{-1}P \otimes p_2^{-1}P) \xrightarrow{\sigma^{-1}\theta} \sigma^{-1}m^{-1}P \to m^{-1}P$$

is the composition of θ with the canonical isomorphism $p_2^{-1}P \otimes p_1^{-1}P \to p_1^{-1}P \otimes p_2^{-1}P$.

▹ *Associativity*: Let $q_i : Q \times Q \times Q \to Q$ be the projection on the i-th factor, where $i \in [\![1, 3]\!]$, and define $q_{ij} : Q \times Q \times Q \to Q \times Q$ similarly, where $(i, j) \in [\![1, 3]\!]^2$ with $i < j$. If $m_3 : Q \times Q \times Q \to Q$ is the multiplication morphism, then the diagram of G-torsors over $Q \times Q \times Q$

$$\begin{array}{ccccc} & & q_1^{-1}P \otimes (mq_{23})^{-1}P & & \\ & \overset{id \otimes q_{23}^{-1}\theta}{\nearrow} & & \overset{(q_1 \times mq_{23})^{-1}\theta}{\searrow} & \\ q_1^{-1}P \otimes q_2^{-1}P \otimes q_3^{-1}P & & & & m_3^{-1}P \\ & \underset{q_{12}^{-1}\theta \otimes id}{\searrow} & & \underset{(mq_{12} \times q_3)^{-1}\theta}{\nearrow} & \\ & & (mq_{12})^{-1}P \otimes q_3^{-1}P. & & \end{array}$$

is commutative.

The category of multiplicative G-torsors is fibered in groupoids over the category of commutative semigroups of E. We denote by $\mathrm{Tors}^{\otimes}(Q, G)$ the groupoid of multiplicative G-torsors over a commutative semigroup Q of E.

Remark 2.6. If $G = \Lambda^\times$ for some ring Λ in E, we use the term "*multiplicative locally free Λ-module of rank* 1" as a synonym for "multiplicative G-torsor", when we want to emphasize the locally free Λ-module of rank 1 corresponding to a given G-torsor, rather than the G-torsor itself (see Example 2.3).

Proposition 2.7. *Let G be an abelian group in E, let Q be a commutative semigroup in E and let I be an ideal of Q. If the projection morphisms $Q \times I \to Q$ and $I \times I \to I$ onto the first factors are morphisms of descent for the fibered category of multiplicative G-torsors* (see Definition 2.5), *then the restriction functor*

$$\mathrm{Tors}^{\otimes}(Q, G) \to \mathrm{Tors}^{\otimes}(I, G)$$

is fully faithful.

Let $i : I \to Q$ be the canonical injection morphism. Let p_1 and p_2 be the projection morphisms of $Q \times I$ onto its first and second factors respectively, and let $m : Q \times I \to I$ be the multiplication morphism. Let (P, θ) and (P', θ') be multiplicative G-torsors over Q. We have an isomorphism

$$\beta_P : p_1^{-1} P \xrightarrow{(\mathrm{id} \times i)^{-1}\theta} m^{-1} i^{-1} P \otimes p_2^{-1} i^{-1} P^{-1},$$

and similarly for P'. If $\alpha : i^{-1}P \to i^{-1}P'$ is a morphism of multiplicative G-torsors over I, then $\beta_{P'}^{-1}(m^{-1}\alpha \otimes p_2^{-1}\alpha)\beta_P$ is an isomorphism from $p_1^{-1}P$ to $p_1^{-1}P'$, which is compatible with the canonical descent datum for p_1 associated to $p_1^{-1}P$ and $p_1^{-1}P'$. Since p_1 is a morphism of descent for the fibered category of multiplicative G-torsors, there is a unique morphism $\gamma : P \to P'$ of multiplicative G-torsors over Q such that $p_1^{-1}\gamma = \beta_{P'}^{-1}(m^{-1}\alpha \otimes p_2^{-1}\alpha)\beta_P$. The restriction of $p_1^{-1}\gamma$ to $I \times I$ is the pullback of α by the first projection, which is a morphism of descent for the fibered category of multiplicative G-torsors, so that the restriction of γ to I is α.

Proposition 2.8. *Let G be an abelian group in E, and let $\rho : M \to Q$ be a morphism of commutative semigroups in E. If ρ (resp. $\rho \times \rho$ and $\rho \times \rho \times \rho$) is a morphism of effective descent (resp. of descent) for the fibered category of G-torsors, then ρ is a morphism of effective descent for the fibered category of multiplicative G-torsors.*

A descent datum of multiplicative G-torsors for ρ yields a descent datum of G-torsors for ρ, hence a G-torsor over Q by hypothesis. Since $\rho \times \rho$ and $\rho \times \rho \times \rho$ are morphisms of descent for the fibered category of G-torsors, the structure of multiplicative G-torsor descends as well. Details are omitted.

Proposition 2.9. *Let G and Q be abelian groups in E. The groupoid* $\mathrm{Tors}^{\otimes}(Q, G)$ *of multiplicative G-torsors over Q is equivalent as a monoidal category to the groupoid of extensions of Q by G in E, with the Baer sum as a monoidal structure.*

We have already seen how to associate a multiplicative G-torsor to an extension of Q by G. This construction is functorial, and the corresponding functor is an equivalence by [Moret-Bailly 1985, I.2.3.10].

Corollary 2.10. *Let G and Q be abelian groups in E. The group of isomorphism classes of multiplicative G-torsors over Q is isomorphic to the group* $\mathrm{Ext}^1(Q, G)$ *of isomorphism classes of extensions of Q by G in E.*

2.11. Let S be a scheme, let X be an S-scheme, and let G be a finite abelian group. Let P be a G-torsor over X in $S_{\text{Ét}}$. Since $P \to X$ is an epimorphism in $S_{\text{Ét}}$, there is an étale cover $(X_i \to X)_{i \in I}$ such that for each $i \in I$, the morphism $X_i \to X$ factors through $P \to X$. In particular, for each $i \in I$ the G-torsor $P \times_X X_i \to X_i$ is isomorphic to the trivial G-torsor $G \times X_i \to X_i$, so that $P \times_X X_i$ is representable by a finite étale X_i-scheme. By étale descent of affine morphisms, we obtain:

Proposition 2.12. *Let G be a finite abelian group, let S be a scheme, and let P be a G-torsor over an S-scheme X in $S_{\text{Ét}}$. Then the étale sheaf $P : \mathrm{Sch}_{/S} \to \mathrm{Sets}$ is representable by a finite étale X-scheme.*

The topos $(S_{\text{Ét}})_{/X}$ coincides with $X_{\text{Ét}}$. The category of G-torsors over X in $S_{\text{Ét}}$ is therefore equivalent to the category of G-torsors over the terminal object in $X_{\text{Ét}}$, and Proposition 2.12 yields:

Corollary 2.13. *Let G be a finite abelian group, let S be a scheme, and let X be an S-scheme. Then the category of G-torsors over X in $S_{\text{Ét}}$ is equivalent to the category of G-torsors over the terminal object in $X_{ét}$.*

2.14. Let S be a scheme, and let G be a finite abelian group. Let Q be a commutative S-group scheme, and let M be a sub-S-semigroup scheme of Q.

Proposition 2.15. *Assume that the morphism*

$$\rho : M \times_S M \to Q, \quad (x, y) \mapsto xy^{-1}$$

is faithfully flat and quasicompact, and that M is flat over S. Then the restriction functor

$$\mathrm{Tors}^{\otimes}(Q, G) \to \mathrm{Tors}^{\otimes}(M, G)$$

is an equivalence of categories.

Let (P, θ) be a multiplicative G-torsor over M. For $i \in [\![1, 4]\!]$, let r_i be the projection of $R = (M \times_S M) \times_{\rho, Q, \rho} (M \times_S M)$ onto its i-th factor. Similarly, for $i, j \in [\![1, 4]\!]$ such that $i < j$, we denote by $r_{ij} : R \to M \times_S M$ the projection on the i-th and j-th factors. We then have a sequence of isomorphisms

$$(r_1^{-1}P \otimes r_2^{-1}P^{-1}) \otimes (r_3^{-1}P \otimes r_4^{-1}P^{-1})^{-1} \longrightarrow r_{14}^{-1}(p_1^{-1}P \otimes p_2^{-1}P) \otimes r_{23}^{-1}(p_1^{-1}P \otimes p_2^{-1}P)^{-1}$$
$$\xrightarrow{r_{14}^{-1}\theta \otimes (r_{23}^{-1}\theta)^{-1}} (mr_{14})^{-1}P \otimes ((mr_{23})^{-1}P)^{-1},$$

of G-torsors over R, where $m : M \times_S M \to M$ is the multiplication of M. Since $mr_{14} = mr_{23}$, the latter G-torsor is canonically trivial. We thus obtain an isomorphism

$$\psi : r_1^{-1}P \otimes r_2^{-1}P^{-1} \to r_3^{-1}P \otimes r_4^{-1}P^{-1},$$

of G-torsors over R. The associativity of θ (see Definition 2.5) implies that ψ is a cocycle, i.e., $(p_1^{-1}P \otimes p_2^{-1}P^{-1}, \psi)$ is a descent datum for ρ. By Proposition 2.12 and since faithfully flat and quasicompact morphisms of schemes are of effective descent for the fibered category of affine morphisms, the conditions of Proposition 2.8 are satisfied, and thus there exists a multiplicative G-torsor P' over Q and an isomorphism $\alpha : \rho^{-1}P' \to p_1^{-1}P \otimes p_2^{-1}P^{-1}$ such that ψ is given by the composition

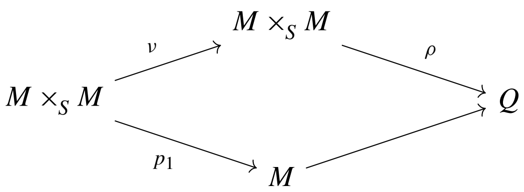
$$r_1^{-1}P \otimes r_2^{-1}P^{-1} \xrightarrow{r_{12}^{-1}\alpha^{-1}} (\rho r_{12})^{-1}P' = (\rho r_{34})^{-1}P' \xrightarrow{r_{34}^{-1}\alpha} r_3^{-1}P \otimes r_4^{-1}P^{-1}.$$

The association $P \mapsto P'$ then defines a functor from $\mathrm{Tors}^{\otimes}(M, G)$ to $\mathrm{Tors}^{\otimes}(Q, G)$. For any multiplicative G-torsor U over Q, we have an isomorphism $U \to (U \times_Q M)'$ by multiplicativity, which is functorial in U.

We now construct, for any multiplicative G-torsor (P, θ) over M, an isomorphism $P \to P' \times_Q M$ of multiplicative G-torsors which is functorial in P. Let $\nu : M \times_S M \to M \times_S M$ be the morphism which sends a section (x, y) to (xy, y). We have an isomorphism

$$(\rho\nu)^{-1}P' \xrightarrow{\nu^{-1}\alpha} \nu^{-1}(p_1^{-1}P \otimes p_2^{-1}P^{-1}) \to m^{-1}P \otimes p_2^{-1}P^{-1} \xrightarrow{\theta^{-1}} p_1^{-1}P.$$

The diagram



is commutative; hence $(\rho\nu)^{-1}P'$ is isomorphic to $p_1^{-1}(P' \times_Q M)$. We thus obtain an isomorphism

$$\beta : p_1^{-1}P \to p_1^{-1}(P' \times_Q M),$$

of multiplicative G-torsors. The morphism β is compatible with the canonical descent data for p_1 associated to $p_1^{-1}P$ and $p_1^{-1}(P' \times_Q M)$. Since p_1 is a covering for the fpqc topology, Proposition 2.8 applies, hence there is a unique isomorphism $\gamma : P \to P' \times_Q M$ of multiplicative G-torsors such that $\beta = p_1^{-1}\gamma$. The construction of this isomorphism of multiplicative G-torsors is functorial in P, hence the result.

2.16. Let A be a ring. If M is an A-module, we denote by $\underline{M}$ the functor $B \mapsto M \otimes_A B$ from Alg_A to Sets.

Definition 2.17 [SGA 4₃ 1973, XVII 5.5.2.2]. Let M and N be A-modules. A *polynomial map* from M to N is a morphism of functors $\underline{M} \to \underline{N}$. A polynomial map $f : \underline{M} \to \underline{N}$ is *homogeneous of degree d* if for any A-algebra B, any element λ of B and any element m of $\underline{M}(B)$, we have $f(\lambda m) = \lambda^d f(m)$.

For each integer d and any A-module M, let $\mathrm{TS}_A^d(M) = (M^{\otimes_A d})^{\mathfrak{S}_d}$ be the A-module of symmetric tensors of degree d with coefficients in M. If M is a free A-module with basis $(e_i)_{i \in I}$, then we have a decomposition

$$TS_A^d(M) = \Bigg(\bigoplus_{\beta : [\![1,d]\!] \to I} A e_{\beta(1)} \otimes \cdots \otimes e_{\beta(d)} \Bigg)^{\mathfrak{S}_d} = \bigoplus_{\substack{\alpha : I \to \mathbb{N} \\ \sum_{i \in I} \alpha(i) = d}} A e_\alpha, \qquad (2.17.1)$$

where we have set

$$e_\alpha = \sum_{\substack{\beta : [\![1,d]\!] \to I \\ \forall i, |\beta^{-1}(\{i\})| = \alpha(i)}} e_{\beta(1)} \otimes \cdots \otimes e_{\beta(d)}.$$

In particular $TS_A^d(M)$ is a free A-module, and its formation commutes with base change by any ring morphism $A \to B$.

Proposition 2.18. *Let M be a flat A-module and let $d \geq 0$ be an integer. Then $\mathrm{TS}_A^d(M)$ is a flat module, and for any A-algebra B the canonical homomorphism*

$$\mathrm{TS}_A^d(M) \otimes_A B \to \mathrm{TS}_B^d(M \otimes_A B)$$

is bijective.

Any flat A-module is a filtered colimit of finite free modules. We have already seen that the conclusion of Proposition 2.18 holds whenever M is free, hence the conclusion in general since the functor TS^d_A commutes with filtered colimits.

Proposition 2.19. *Let M be a flat A-module and let $d \geq 0$ be an integer. Let $\gamma_d : \underline{M} \to \underline{\mathrm{TS}^d_A(C)}$ be the functor which sends, for any A-algebra B, an element m of $\underline{M}(B)$ to the element $m^{\otimes d}$ of $\mathrm{TS}^d_B(M \otimes_A B) = \mathrm{TS}^d_A(M) \otimes_A B$* (see Proposition 2.18). *Then, for any homogeneous polynomial map $f : \underline{M} \to \underline{N}$ of degree d, there is a unique A-linear homomorphism $\tilde{f} : \mathrm{TS}^d_A(M) \to N$ such that $f = \tilde{f}\gamma_d$.*

As in Proposition 2.18, we can assume that M is free of finite rank over A. Let $(e_i)_{i\in I}$ be a basis of M. Let us write

$$f\Big(\sum_{i\in I} X_i e_i\Big) = \sum_{\alpha : I \to \mathbb{N}} X^\alpha f_\alpha$$

in $\underline{N}(A[(X_i)_{i\in I}])$ for some elements $(f_\alpha)_\alpha$ of N, where $X^\alpha = \prod_{i\in I} X_i^{\alpha_i}$. Accordingly, we have for any A-algebra B and any element $m = \sum_{i\in I} b_i e_i$ of $\underline{M}(B)$, the formula

$$f(m) = \sum_{\alpha : I \to \mathbb{N}} b^\alpha f_\alpha,$$

where $b^\alpha = \prod_{i\in I} b_i^{\alpha_i}$, by using the naturality of f with the unique morphism of A-algebras $A[(X_i)_{i\in I}] \to B$ which sends X_i to b_i for each i. By applying this to the element $m = \sum_{i\in I} T X_i e_i$ of $\underline{M}(A[T, (X_i)_{i\in I}])$, we obtain

$$f\Big(\sum_{i\in I} T X_i e_i\Big) = \sum_{\alpha : I \to \mathbb{N}} T^{|\alpha|} X^\alpha f_\alpha,$$

where we have set $|\alpha| = \sum_{i\in I} \alpha(i)$. Since f is homogeneous of degree d, the left side of this equation is also equal to

$$T^d f\Big(\sum_{i\in I} X_i e_i\Big) = \sum_{\alpha : I \to \mathbb{N}} T^d X^\alpha f_\alpha.$$

We conclude that $T^d f_\alpha = T^{|\alpha|} f_\alpha$ in $N \otimes_A A[T]$ for any $\alpha : I \to \mathbb{N}$, and thus that $f_\alpha = 0$ whenever $|\alpha|$ differs from d. We therefore have

$$f(m) = \sum_{\substack{\alpha : I \to \mathbb{N} \\ |\alpha| = d}} b^\alpha f_\alpha,$$

for any A-algebra B and any element $m = \sum_{i\in I} b_i e_i$ of $\underline{M}(B)$. Using the decomposition (2.17.1), we also have

$$\gamma_d(m) = \sum_{\beta : [\![1,d]\!] \to I} \otimes_{j=1}^d b_{\beta(j)} e_{\beta(j)} = \sum_{\substack{\alpha : I \to \mathbb{N} \\ |\alpha| = d}} b^\alpha e_\alpha.$$

The conclusion of Proposition 2.19 is achieved by taking $\tilde{f}$ to be the unique morphism of A-modules from $\mathrm{TS}^d_A(M)$ to N which sends e_α to f_α.

2.20. Let $A \to C$ be a ring morphism such that C is a finitely generated projective A-module of rank d. For any A-algebra B and any element m of $\underline{C}(B)$, we set

$$N_{C/A}(c) = \det_{\underline{A}(B)}(m_c),$$

where m_c is the $\underline{A}(B)$-linear endomorphism of $\underline{C}(B)$ induced by the multiplication by c. This defines a homogeneous polynomial map $N_{C/A} : \underline{C} \to \underline{A}$ of degree d (see Definition 2.17). By Proposition 2.19, there is a unique morphism of A-modules $\varphi : \mathrm{TS}^d_A(C) \to A$ such that $N_{C/A} = \varphi\gamma_d$.

Proposition 2.21 [SGA 4_3 1973, XVII 6.3.1.6]. *The morphism of A-modules $\varphi : \mathrm{TS}^d_A(C) \to A$ is a morphism of A-algebras.*

Let x be an element of C, and let us consider the morphism of A-modules $f : y \to \varphi(\gamma_d(x)y)$ from $\mathrm{TS}^d_A(C)$ to A. For any A-algebra B and any element c of $\underline{C}(B)$, we have

$$f(\gamma_d(c)) = \varphi(\gamma_d(x)\gamma_d(c)) = \varphi(\gamma_d(xc)) = N_{C/A}(xc) = N_{C/A}(x)N_{C/A}(c)$$

by the multiplicativity of determinants, so that $f(\gamma_d(c)) = N_{C/A}(x)\varphi(\gamma_d(c))$. By the uniqueness statement in Proposition 2.19, we obtain $f = N_{C/A}(x)\varphi$, i.e., for all y in $\mathrm{TS}^d_A(C)$ we have

$$\varphi(\gamma_d(x)y) = N_{C/A}(x)\varphi(y). \tag{2.21.1}$$

For any A-algebra B, one can apply this argument to the morphism $B \to \underline{C}(B)$ instead of $A \to C$. Thus (2.21.1) also holds for any element x of $\underline{C}(B)$ and any element y of $\underline{\mathrm{TS}^d_A(C)}(B) = \mathrm{TS}^d_{\underline{A}(B)}(\underline{C}(B))$ (see Proposition 2.18). Now, let y be an element of $\mathrm{TS}^d_A(C)$ and let us consider the morphism of A-modules $g : z \to \varphi(zy)$ from $\mathrm{TS}^d_A(C)$ to A. We have proved that $g\gamma_d = \varphi(y)N_{C/A}$, hence $g = \varphi(y)\varphi$ by Proposition 2.19. Thus φ is a morphism of rings. Since φ is also A-linear, it is a morphism of A-algebras.

2.22. Let S be a scheme.

Definition 2.23 [SGA 1 1971, V.1.7].

- ▹ Let T be an object of a category C endowed with a right action of a group Γ. We say that *the quotient T/Γ exists* in C if the covariant functor

  $$C \to \mathrm{Sets}, \quad U \mapsto \mathrm{Hom}_C(T, U)^\Gamma$$

  is representable by an object of C.

- ▹ Let T be an S-scheme. An action of a finite group Γ on T is *admissible* if there exists an affine Γ-invariant morphism $f : T \to T'$ such that the canonical morphism $\mathcal{O}_{T'} \to f_*\mathcal{O}_T$ induces an isomorphism from $\mathcal{O}_{T'}$ to $(f_*\mathcal{O}_T)^\Gamma$.

Proposition 2.24 [SGA 1 1971, V.1.3]. *Let T be an S-scheme endowed with an admissible right action of a finite group Γ. If $f : T \to T'$ is an affine Γ-invariant morphism such that the canonical morphism $\mathcal{O}_{T'} \to f_*\mathcal{O}_T$ induces an isomorphism from $\mathcal{O}_{T'}$ to $(f_*\mathcal{O}_T)^\Gamma$, then the quotient T/Γ exists and is isomorphic to T'.*

Proposition 2.25 [SGA 1 1971, V.1.8]. *Let T be an S-scheme endowed with a right action of a finite group Γ. The action of Γ on T is admissible if and only if T is covered by Γ-invariant affine open subsets.*

Proposition 2.26 [SGA 1 1971, V.1.9]. *Let T be an S-scheme endowed with an admissible right action of a finite group Γ, and let S' be a flat S-scheme. Then, the action of Γ on the S'-scheme $T \times_S S'$ is admissible, and the canonical morphism*

$$(T \times_S S')/\Gamma \to (T/\Gamma) \times_S S'$$

is an isomorphism.

Let X be an S-scheme and let $d \geq 0$ be an integer. The group $\mathfrak{S}_d$ of permutations of $[\![1, d]\!]$ acts on the right on the S-scheme $X^{\times_S d} = X \times_S \cdots \times_S X$ by the formula

$$(x_i)_{i \in [\![1,d]\!]} \cdot \sigma = (x_{\sigma(i)})_{i \in [\![1,d]\!]}.$$

Proposition 2.27. *If X is Zariski-locally quasiprojective over S, then the right action of $\mathfrak{S}_d$ on $X^{\times_S d}$ is admissible. In particular, the quotient $\mathrm{Sym}^d_S(X) = X^{\times_S d}/\mathfrak{S}_d$ exists in the category of S-schemes.*

Since X is Zariski-locally quasiprojective over S, any finite set of points in X with the same image in S is contained in an affine open subset of X. Thus $X^{\times_S d}$ is covered by open subsets of the form $U^{\times_S d}$ where U is an affine open subset of X whose image in S is contained in an affine open subset of S. These particular open subsets are affine and $\mathfrak{S}_d$-invariant, so that the action of $\mathfrak{S}_d$ on $X^{\times_S d}$ is admissible by Proposition 2.25.

Remark 2.28. If $X = \mathrm{Spec}(B)$ and $S = \mathrm{Spec}(A)$ are affine, then for any S-scheme T we have

$$\mathrm{Hom}_{\mathrm{Sch}_{/S}}(X^{\times_S d}, T)^{\mathfrak{S}_d} = \mathrm{Hom}_{\mathrm{Alg}_A}(\Gamma(T, \mathcal{O}_T), B^{\otimes_A d})^{\mathfrak{S}_d} = \mathrm{Hom}_{\mathrm{Alg}_A}(\Gamma(T, \mathcal{O}_T), \mathrm{TS}^d_A(B)),$$

see Section 2.16. Thus $\mathrm{Sym}^d_S(X)$ is representable by the S-scheme $\mathrm{Spec}(\mathrm{TS}^d_A(B))$.

Proposition 2.29. *If X is flat and Zariski-locally quasiprojective over S, then $\mathrm{Sym}^d_S(X)$ is flat over S. Moreover, for any S-scheme S', the canonical morphism*

$$\mathrm{Sym}^d_{S'}(X \times_S S') \to \mathrm{Sym}^d_S(X) \times_S S'$$

is an isomorphism.

This follows from Remark 2.28 and from Proposition 2.18.

Proposition 2.30 [SGA 1 1971, IX.5.8]. *Let G be a finite abelian group, let P be a G-torsor over an S-scheme X in $S_{\text{Ét}}$. Assume that P and X are endowed with right actions from a finite group Γ such that the morphism $P \to X$ is Γ-equivariant, and that the following properties hold:*

(a) *The right Γ-action on P commutes with the left G-action.*

(b) *The right Γ-action on X is admissible* (see Definition 2.23), *and the quotient morphism $X \to X/\Gamma$ is finite.*

(c) *For any geometric point $\bar{x}$ of X, the action of the stabilizer $\Gamma_{\bar{x}}$ of $\bar{x}$ in Γ on the fiber $P_{\bar{x}}$ of P at $\bar{x}$ is trivial.*

Then the action of Γ on P is admissible, and P/Γ is a G-torsor over X/Γ in $S_{\text{Ét}}$.

2.31. Let S be a scheme, let X be an S-scheme and let $d \geq 1$ be an integer. Let G be a finite abelian group, and let $P \to X$ be a G-torsor over X in $S_{\text{Ét}}$. By Proposition 2.12, the sheaf P is representable by a finite étale X-scheme.

For each $i \in [\![1, d]\!]$ let $p_i : X^{\times_S d} \to X$ be the projection on i-th factor, and let us consider the G-torsor

$$p_1^{-1} P \otimes \cdots \otimes p_d^{-1} P = G_d \setminus P^{\times_S d}$$

over $X^{\times_S d}$, where $G_d \subseteq G^d$ is the kernel of the multiplication morphism $G^d \to G$. By Proposition 2.12, the object $G_d \setminus P^{\times_S d}$ of $S_{\text{Ét}}$ is representable by an S-scheme which is finite étale over $X^{\times_S d}$. The group $\mathfrak{S}_d$ acts on the right on $G_d \setminus P^{\times_S d}$ by the formula

$$(p_i)_{i \in [\![1,d]\!]} \cdot \sigma = (p_{\sigma(i)})_{i \in [\![1,d]\!]}.$$

This action of $\mathfrak{S}_d$ commutes with the left action of G on $G_d \setminus P^{\times_S d}$.

Proposition 2.32. *If X is Zariski-locally quasiprojective on S, then the right action of $\mathfrak{S}_d$ on $G_d \setminus P^{\times_S d}$ is admissible* (see Definition 2.23), *so that the quotient $P^{[d]}$ of $G_d \setminus P^{\times_S d}$ by $\mathfrak{S}_d$ exists in* $\mathrm{Sch}_{/S}$. *Moreover, the canonical morphism $P^{[d]} \to \mathrm{Sym}_S^d(X)$ is a G-torsor, and the morphism*

$$p_1^{-1} P \otimes \cdots \otimes p_d^{-1} P \to r^{-1} P^{[d]}$$

where $r : X^{\times_S d} \to \mathrm{Sym}_S^d(X)$ is the canonical projection, is an isomorphism of G-torsors over $X^{\times_S d}$.

By Propositions 2.27 and 2.30, it is sufficient to show that if $\bar{x} = (\bar{x}_i)_{i=1}^d$ is a geometric point of $X^{\times_S d}$, then the stabilizer of $\bar{x}$ in $\mathfrak{S}_d$ acts trivially on $(G_d \setminus P^{\times_S d})_{\bar{x}}$. Assume that the finite set $\{\bar{x}_i \mid i \in [\![1, d]\!]\}$ has exactly r distinct elements $\bar{y}_1, \ldots, \bar{y}_r$, where $\bar{y}_j$ appears with multiplicity d_j. Then the stabilizer of $\bar{x}$ in $\mathfrak{S}_d$ is isomorphic to the subgroup $\prod_{j=1}^r \mathfrak{S}_{d_j}$ of $\mathfrak{S}_d$. For each $j \in [\![1, r]\!]$, the G-torsor $P_{\bar{y}_j}$ is trivial, and if e is a section of this torsor then $(e)_{i=1}^{d_j}$ is a section of $G_{d_j} \setminus P_{\bar{y}_j}^{d_j}$ which is $\mathfrak{S}_{d_j}$-invariant. The action of $\mathfrak{S}_{d_j}$ on $G_{d_j} \setminus P_{\bar{y}_j}^{d_j}$ is therefore trivial, so that the action of $\prod_{j=1}^r \mathfrak{S}_{d_j}$ on the G-torsor

$$(G_d \setminus P^{\times_S d})_{\bar{x}} = G_r \setminus \left(\prod_{j=1}^r G_{d_j} \setminus P_{\bar{y}_j}^{d_j} \right)$$

is trivial as well.

Proposition 2.33. *If X is flat and Zariski-locally quasiprojective on S, then for any S-scheme S', the canonical morphism*

$$(P \times_S S')^{[d]} \to P^{[d]} \times_S S'$$

is an isomorphism.

By Proposition 2.29, the canonical morphism

$$\mathrm{Sym}^d_{S'}(X \times_S S') \to \mathrm{Sym}^d_S(X) \times_S S'$$

is an isomorphism. Thus the second morphism in the composition

$$(P \times_S S')^{[d]} \to (P^{[d]} \times_S S') \times_{\mathrm{Sym}^d_S(X) \times_S S'} \mathrm{Sym}^d_{S'}(X \times_S S') \to P^{[d]} \times_S S'$$

is an isomorphism, while the first morphism is a morphism of G-torsors, hence an isomorphism.

3. Geometric local class field theory

Let k be a perfect field, and let L be a complete discretely valued extension of k with residue field k. We denote by $\mathcal{O}_L$ its ring of integers, and by $\mathfrak{m}_L$ the maximal ideal of $\mathcal{O}_L$.

3.1. Let us consider the functor

$$\mathbb{O}_L : \mathrm{Alg}_k \to \mathrm{Alg}_{\mathcal{O}_L}, \quad A \mapsto \lim_n A \otimes_k \mathcal{O}_L/\mathfrak{m}_L^n,$$

with values in the category of $\mathcal{O}_L$-algebras.

Proposition 3.2. *The functor $\mathbb{O}_L$ is representable by a k-scheme.*

Indeed, if π is a uniformizer of L, then we have an isomorphism $k((t)) \to L$ which sends t to π, so that the functor $\mathbb{O}_L$ is isomorphic to the functor $A \mapsto A[[t]]$, which is representable by an affine space over k of countable dimension.

Corollary 3.3. *The functor $\mathbb{L} = \mathbb{O}_L \otimes_{\mathcal{O}_L} L$ is representable by an ind-k-scheme.*

We can assume that L is the field of Laurent series $k((t))$. In this case, we have

$$\mathbb{L}(A) = A((t)) = \mathrm{colim}_n\, t^{-n} A[[t]]$$

for any k-algebra A, and for each integer n the functor $A \mapsto t^{-n}A[[t]]$ is representable by a k-scheme, see Proposition 3.2.

Proposition 3.4. *Let G (resp. H) be the functor from Alg_k to the category of groups which associates to a k-algebra A the subgroup $G(A)$ of $A((t))^\times$ consisting of Laurent series of the form $1 + \sum_{r>0} a_r t^{-r}$ where a_r is a nilpotent element of A for each $r > 0$ and vanishes for r large enough (resp. of Laurent series of the form $1 + \sum_{r>0} a_r t^r$ where a_r belongs to A for each $r > 0$). Let $\underline{\mathbb{Z}}$ be the functor which sends a k-algebra A to the group of locally constant functions $\mathrm{Spec}(A) \to \mathbb{Z}$. For any uniformizer π of L, the morphism*

$$\mathbb{G}_{m,k} \times \underline{\mathbb{Z}} \times G \times H \to \mathbb{L}^\times, \quad (a, n, g, h) \mapsto a\pi^n g(\pi) h(\pi),$$

is an isomorphism of group-valued functors.

Let A be a k-algebra. By [Contou-Carrère 2013, 0.8], every invertible element u of $A((t))$ uniquely factors as $u = t^n f(t)h(t)$ where $f(t)$ and $h(t)$ are elements of $A[[t]]^\times$ and $G(A)$ respectively, and $n : \mathrm{Spec}(A) \to \mathbb{Z}$ is a locally constant function. Moreover, there is a unique factorization $f(t) = ag(t)$ where a and $g(t)$ belong to $A^\times$ and $H(A)$ respectively, hence the result.

Corollary 3.5. *The functor $\mathbb{L}^\times$ is representable by an ind-k-scheme. Moreover, its restriction to the category of reduced k-algebras is representable by a reduced k-scheme.*

The groups $\underline{\mathbb{Z}}$ and H from Proposition 3.4 are representable by reduced k-schemes, and so is $\mathbb{G}_{m,k}$. The group G from Proposition 3.4 is the filtered colimit of the functor $n \mapsto G_n$, where G_n is the functor which associates to a k-algebra A the subset $G_n(A)$ of $A((t))^\times$ consisting of Laurent series of the form $1 + \sum_{r=1}^{n} a_r t^{-r}$ where $a_r^n = 0$ for each $r \in [\![1, n]\!]$. For each n, the functor G_n is representable by an affine k-scheme. Thus G is representable by an ind-k-scheme, and so is $\mathbb{L}^\times$ by Proposition 3.4. The last assertion of Corollary 3.5 follows from the fact that $G(A)$ is the trivial group for any reduced k-algebra A.

Corollary 3.6. *Let $d \geq 0$ be an integer. Let $\mathbb{U}_L^{(d)}$ be the subfunctor of $\mathbb{L}^\times$ given by $1 + \mathfrak{m}_L^d \mathbb{O}_L$ if $d \geq 1$ and by $\mathbb{O}_L^\times$ if $d = 0$. Then the functor*

$$\mathbb{L}^\times/\mathbb{U}_L^{(d)} : \mathrm{Alg}_k \to \mathrm{Sets}, \quad A \mapsto \mathbb{L}^\times(A)/\mathbb{U}_L^{(d)}(A),$$

is representable by an ind-k-scheme. Moreover, its restriction to the category of reduced k-algebras is representable by a reduced k-scheme.

According to Proposition 3.4, it is sufficient to show that $(\mathbb{G}_{m,k} \times H)/\mathbb{U}_{k((t))}^{(d)}$ is representable by a reduced k-scheme. The case $d = 0$ is clear, while for $d \geq 1$, we have for any k-algebra A a bijection

$$A^\times \times A^{[\![1,d-1]\!]} \to (\mathbb{G}_{m,k} \times H)(A)/\mathbb{U}_{k((t))}^{(d)}(A), \quad (a_i)_{0 \leq i \leq d-1} \mapsto \sum_{i=0}^{d-1} a_i t^i;$$

hence the result.

3.7. From now on, we consider $\mathrm{Spec}(L)$, $\mathbb{L}^\times$ and $\mathbb{L}^\times/\mathbb{U}_L^{(d)}$ for each integer $d \geq 0$ as objects of the topos $\mathrm{Spec}(k)_{\text{Ét}}$. Let π be an uniformizer of L. We denote by Π the element of $\mathbb{L}(k)$ corresponding to π via the canonical identification $L \simeq \mathbb{L}(k)$. Thus the functor $\mathbb{L}^\times$ is given by

$$\mathbb{L}^\times : A \in \mathrm{Alg}_k \mapsto A((\Pi))^\times.$$

In particular, the Laurent series $(\Pi - \pi)^{-1}\Pi = -\sum_{n\geq 1} \pi^{-n}\Pi^n$ defines an L-point of $\mathbb{L}^\times$. We denote by $\varphi : \mathrm{Spec}(L) \to \mathbb{L}^\times$ the corresponding morphism. We follow here Contou-Carrère's convention; in [Suzuki 2013], the morphism φ corresponds to the point $(\Pi - \pi)\Pi^{-1}$ instead. This is harmless since the inversion is an automorphism of the abelian group $\mathbb{L}^\times$.

Theorem 3.8 [Suzuki 2013, Theorem A(1)]. *Let G be a finite abelian group. The functor*

$$\mathrm{Tors}^{\otimes}(\mathbb{L}^\times, G) \to \mathrm{Tors}(\mathrm{Spec}(L), G), \quad P \to \varphi^{-1}P,$$

is an equivalence of categories (see Definitions 2.2 and 2.5).

In the case where k is algebraically closed, Serre [1961] constructed an equivalence

$$\mathrm{Tors}(\mathrm{Spec}(L), G) \to \mathrm{Tors}^{\otimes}(\mathbb{L}^{\times}, G).$$

Suzuki [2013] shows that the functor from Theorem 3.8 is a quasiinverse to Serre's functor when k is algebraically closed, and extends the result to arbitrary perfect residue fields. In particular, the equivalence from Theorem 3.8 is canonical, even though its definition depends on the choice of π. Suzuki's proof of Theorem 3.8 relies on the Albanese property of the morphism φ, previously established by Contou-Carrère.

Let L^{sep} be a separable closure of L, and let G_L be the Galois group of L^{sep} over L, so that the small étale topos of $\mathrm{Spec}(L)$ is isomorphic to the topos of sets with continuous left G_L-action. By Corollary 2.13, the category of G-torsors over $\mathrm{Spec}(L)$ in $\mathrm{Spec}(k)_{\text{Ét}}$ is isomorphic to the category of G-torsors in the small étale topos $\mathrm{Spec}(L)_{\text{ét}}$. Correspondingly, for each finite abelian group G, the group of isomorphism classes of the category $\mathrm{Tors}(\mathrm{Spec}(L), G)$ is isomorphic to the group of continuous homomorphisms from G_L to G.

We denote by $(G_L^j)_{j\geq -1}$ the ramification filtration of G_L [Serre 1962, IV.3], so that $G_L^{-1} = G_L$ and G_L^0 is the inertia subgroup of G_L, while $G_L^{0+} = \cup_{j>0} G_L^j$ is the wild inertia subgroup of G_L.

Definition 3.9. Let G be a finite abelian group and let $d \geq 0$ be a rational number. A G-torsor over $\mathrm{Spec}(L)$ (in $\mathrm{Spec}(k)_{\text{Ét}}$), corresponding to a continuous homomorphism $\rho : G_L \to G$, is said to have *ramification bounded by d* if $\rho(G_L^d) = \{1\}$. A G-torsor over $\mathrm{Spec}(L)$ with ramification bounded by 0 or 1 is said to be unramified or tamely ramified, respectively.

Proposition 3.10. *Let G be a finite abelian group, let $d \geq 0$ be an integer, and let P be a multiplicative G-torsor P over $\mathbb{L}^{\times}$* (see Definition 2.5). *Assume that k is algebraically closed. Then $\varphi^{-1}P$ has ramification bounded by d* (see Definition 3.9) *if and only if P is the pullback of a multiplicative G-torsor over $\mathbb{L}^{\times}/\mathbb{U}_L^{(d)}$* (see Corollary 3.6).

This follows from [Serre 1961, 3.2 Theorem 1] and from the compatibility of φ^{-1} with Serre's construction [Suzuki 2013, Theorem A(2)].

3.11. Let π and φ be as in Section 3.7. Let K be a closed subextension of k in L, such that $K \to L$ is a finite extension of degree d. Since L is a finite free K-algebra of rank d, we have a canonical morphism of K-schemes

$$\psi : \mathrm{Spec}(K) \to \mathrm{Sym}_K^d(\mathrm{Spec}(L))$$

by Proposition 2.21.

Proposition 3.12. *The composition*

$$\mathrm{Spec}(K) \xrightarrow{\psi} \mathrm{Sym}_K^d(\mathrm{Spec}(L)) \to \mathrm{Sym}_k^d(\mathrm{Spec}(L)) \xrightarrow{\mathrm{Sym}_k^d(\varphi)} \mathrm{Sym}_k^d(\mathbb{L}^{\times}) \to \mathbb{L}^{\times},$$

where the last morphism is given by the multiplication, corresponds to the K-point $P_\pi(\Pi)^{-1}\Pi^d$ of $\mathbb{L}^{\times}$, where the polynomial P_π is the characteristic polynomial of the K-linear endomorphism $x \mapsto \pi x$ of L.

We first describe the morphism ψ. The scheme $\mathrm{Sym}^d_K(\mathrm{Spec}(L))$ is the spectrum of the k-algebra $\mathrm{TS}^d_K(L)$ of symmetric tensors of degree d in L, see Proposition 2.27. The elements $e_i = \pi^{i-1}$ for $i = 1, \ldots, d$ form a K-basis of L, so that we have a decomposition

$$\mathrm{TS}^d_K(L) = \bigoplus_{\substack{\alpha:[\![1,d]\!]\to\mathbb{N} \\ \sum_i \alpha(i)=d}} K e_\alpha,$$

where we have set (see Section 2.16)

$$e_\alpha = \sum_{\substack{\beta:[\![1,d]\!]\to[\![1,d]\!] \\ \forall i, |\beta^{-1}(\{i\})|=\alpha(i)}} e_{\beta(1)} \otimes \cdots \otimes e_{\beta(d)}.$$

Let us write the norm polynomial as

$$N_{L/K}\left(\sum_{i=1}^d x_i e_i\right) = \sum_{\substack{\alpha:[\![1,d]\!]\to\mathbb{N} \\ \sum_i \alpha(i)=d}} f_\alpha x^\alpha,$$

where $x^\alpha = x_1^{\alpha(1)} \cdots x_d^{\alpha(d)}$, and the f_α are uniquely determined elements of K. The morphism $\mathrm{TS}^d_K(L) \to K$ corresponding to ψ is the unique K-linear homomorphism which sends e_α to f_α (see Proposition 2.19 and its proof).

Next we describe the composition

$$\mathrm{Sym}^d_K(\mathrm{Spec}(L)) \to \mathrm{Sym}^d_k(\mathrm{Spec}(L)) \xrightarrow{\mathrm{Sym}^d_k(\varphi)} \mathrm{Sym}^d_k(\mathbb{L}^\times) \to \mathbb{L}^\times.$$

Its precomposition with the projection $\mathrm{Spec}(L)^{\times_K d} \to \mathrm{Sym}^d_K(\mathrm{Spec}(L))$ corresponds to the element of $L^{\otimes_K d}((\Pi))^\times$ given by the formula

$$\prod_{i=1}^d ((\Pi - 1^{\otimes(i-1)} \otimes \pi \otimes 1^{\otimes(d-i)})^{-1}\Pi) = P(\Pi)^{-1}\Pi^d,$$

where the polynomial $P(\Pi)$ can be computed as follows:

$$P(\Pi) = \prod_{i=1}^d \left(\Pi - 1^{\otimes(i-1)} \otimes \pi \otimes 1^{\otimes(d-i)}\right) = \sum_{r=0}^d (-1)^r \Pi^{d-r} \sum_{\substack{(i_1,\ldots,i_d)\in\{0,1\}^d \\ |\{s|i_s=1\}|=r}} \pi^{i_1} \otimes \cdots \otimes \pi^{i_d} = \sum_{r=0}^d (-1)^r e_{\alpha_r} \Pi^{d-r},$$

where $\alpha_r : [\![1, d]\!] \to \mathbb{N}$ is the map which sends 1 and 2 to $d - r$ and r respectively, and any $i > 2$ to 0. The image of $P(\Pi)$ by ψ in $K[\Pi]$ is the polynomial

$$\sum_{r=0}^d (-1)^r f_{\alpha_r} \Pi^{d-r} = N_{L[\Pi]/K[\Pi]}(\Pi e_1 - e_2).$$

Since $e_1 = 1$ and $e_2 = \pi$, we obtain Proposition 3.12.

Proposition 3.13. *Let G be a finite abelian group, and let Q be a G-torsor over* $\mathrm{Spec}(L)$ *(in* $\mathrm{Spec}(k)_{\text{Ét}}$*) of ramification bounded by d* (see Definition 3.9)*. Then $\psi^{-1}Q^{[d]}$* (see Proposition 2.32) *is tamely ramified on* $\mathrm{Spec}(K)$.

Let K' be the maximal unramified extension of K in a separable closure of K. The formation of $\mathrm{Sym}^d_K(\mathrm{Spec}(L))$ is compatible with any base change by Proposition 2.26 or by Proposition 2.29, and so is the formation of φ. Moreover, a G-torsor over $\mathrm{Spec}(K)$ is tamely ramified if and only if its restriction to $\mathrm{Spec}(K')$ is tamely ramified. By replacing K and L by K' and the components of $K'\otimes_K L$ respectively, we can assume that the residue field k is algebraically closed.

Let P be the multiplicative G-torsor on $\mathbb{L}^\times$ (see Definition 2.5) associated to Q (see Theorem 3.8), so that Q is isomorphic to $\varphi^{-1}P$. Then $\psi^{-1}Q^{[d]}$ is isomorphic to the pullback of P along the composition

$$\mathrm{Spec}(K)\xrightarrow{\psi}\mathrm{Sym}^d_K(\mathrm{Spec}(L))\to\mathrm{Sym}^d_k(\mathrm{Spec}(L))\xrightarrow{\mathrm{Sym}^d_k(\varphi)}\mathrm{Sym}^d_k(\mathbb{L}^\times)\to\mathbb{L}^\times$$

considered in Proposition 3.12. By Proposition 3.12, this composition corresponds to the K-point of $\mathbb{L}^\times$ given by $P_\pi(\Pi)^{-1}\Pi^d$, where P_π is the characteristic polynomial of π acting K-linearly by multiplication on L. Let us consider the morphism of pointed sets

$$\begin{aligned}\rho:\mathbb{L}^\times(K)&\to H^1(\mathrm{Spec}(K)_{\text{Ét}},G)\\ \nu&\to\nu^{-1}P\end{aligned}$$

where an element ν of $\mathbb{L}^\times(K)$ is identified to a morphism $\mathrm{Spec}(K)\to\mathbb{L}^\times$. If ν_1 and ν_2 are elements of $\mathbb{L}^\times(K)$, then using the isomorphism $\theta: p_1^{-1}P\otimes p_2^{-1}P\to m^{-1}P$ from Definition 2.5, we obtain isomorphisms

$$(\nu_1\nu_2)^{-1}P\leftarrow(\nu_1\times\nu_2)^{-1}m^{-1}P\xleftarrow{(\nu_1\times\nu_2)^{-1}\theta}(\nu_1\times\nu_2)^{-1}(p_1^{-1}P\otimes p_2^{-1}P)\leftarrow\nu_1^{-1}P\otimes\nu_2^{-1}P.$$

Thus ρ is an homomorphism of abelian groups.

We have to prove that $\rho(\nu)$ is the isomorphism class of a tamely ramified G-torsor over $\mathrm{Spec}(K)$, where $\nu=P_\pi(\Pi)^{-1}\Pi^d$. Since P_π is an Eisenstein polynomial, it can be written as $P_\pi(\Pi)=\Pi^d+cR(\Pi)$, where $c=P_\pi(0)$ is a uniformizer of K, and R is a polynomial of degree $<d$ with coefficients in $\mathcal{O}_K$, such that $R(0)=1$. Thus we can write

$$\nu=c^{-1}\nu_1\nu_2,$$

where $\nu_1=R(\Pi)^{-1}\Pi^d$ and $\nu_2=(1+c^{-1}\Pi^dR(\Pi)^{-1})^{-1}$, so that $\rho(\nu)=\rho(c)^{-1}\rho(\nu_1)\rho(\nu_2)$.

Since Q has ramification bounded by d (see Definition 3.9), the restriction of ρ to $\mathbb{U}_L^{(d)}(K)$ is trivial (see Proposition 3.10). In particular, $\rho(\nu_2)$ is trivial since ν_2 belongs to $\mathbb{U}_L^{(d)}(K)$.

The element ν_1 belongs to $\mathbb{L}^\times(\mathcal{O}_K)$, so that the morphism $\nu_1:\mathrm{Spec}(K)\to\mathbb{L}^\times$ factors through $\mathrm{Spec}(\mathcal{O}_K)$. This implies that $\rho(\nu_1)$ is the isomorphism class of an unramified G-torsor over $\mathrm{Spec}(K)$. It remains to prove that $\rho(c)$ is the isomorphism class of a tamely ramified G-torsor over $\mathrm{Spec}(K)$. Since c belongs to $K^\times=\mathbb{G}_{m,k}(K)\subseteq\mathbb{L}^\times(K)$, this is a consequence of the following lemma:

Lemma 3.14. *Let T be a multiplicative G-torsor over the k-group scheme $\mathbb{G}_{m,k}$* (see Definition 2.5). *Then T is tamely ramified at* 0 *and* ∞.

Let G_k be the constant k-group scheme associated to k. By Proposition 2.9, there is a structure of k-group scheme on T and an exact sequence

$$1 \to G_k \to T \to \mathbb{G}_{m,k} \to 1 \tag{3.14.1}$$

in $\mathrm{Spec}(k)_{\text{Ét}}$, such that the structure of G-torsor on T is given by the action of its subgroup G by translations. Since the fppf topology is finer than the étale topology on $\mathrm{Sch}_{/k}$, the sequence (3.14.1) remains exact in the topos $\mathrm{Spec}(k)_{\mathrm{Fppf}}$. In particular, we obtain a class in the group $\mathrm{Ext}^1_{\mathrm{Fppf}}(\mathbb{G}_{m,k}, G_k)$ of extensions of $\mathbb{G}_{m,k}$ by G_k in $\mathrm{Spec}(k)_{\mathrm{Fppf}}$.

Let $n = |G|$. In the topos $\mathrm{Spec}(k)_{\mathrm{Fppf}}$ we have an exact sequence

$$1 \to \mu_{n,k} \to \mathbb{G}_{m,k} \xrightarrow{n} \mathbb{G}_{m,k} \to 1, \tag{3.14.2}$$

where $\mu_{n,k}$ is the k-group scheme of n-th roots of unity. By applying the functor $\mathrm{Hom}(\cdot, G_k)$, we obtain an exact sequence

$$\mathrm{Hom}(\mu_{n,k}, G_k) \xrightarrow{\delta} \mathrm{Ext}^1_{\mathrm{fppf}}(\mathbb{G}_{m,k}, G_k) \xrightarrow{n} \mathrm{Ext}^1_{\mathrm{fppf}}(\mathbb{G}_{m,k}, G_k).$$

Since $n = |G|$, the group $\mathrm{Ext}^1_{\mathrm{Fppf}}(\mathbb{G}_{m,k}, G_k)$ is annihilated by n, so that the homomorphism δ above is surjective. Thus the exact sequence (3.14.1) in $\mathrm{Spec}(k)_{\mathrm{Fppf}}$ is the pushout of (3.14.2) along an homomorphism $\mu_{n,k} \to G_k$. Let n' be the largest divisor of n which is invertible in k. Then the largest étale quotient of $\mu_{n,k}$ is the epimorphism $\mu_{n,k} \to \mu_{n',k}$ given by $x \mapsto x^{n/n'}$. In particular, the homomorphism $\mu_{n,k} \to G_k$ factors through $\mu_{n',k}$, so that (3.14.1) is the pushout of the extension

$$1 \to \mu_{n',k} \to \mathbb{G}_{m,k} \xrightarrow{n'} \mathbb{G}_{m,k} \to 1$$

along an homomorphism $\mu_{n',k} \to G_k$. Since the morphism $\mathbb{G}_{m,k} \xrightarrow{n'} \mathbb{G}_{m,k}$ is tamely ramified above 0 and ∞, so is the morphism $T \to \mathbb{G}_{m,k}$.

4. Rigidified Picard schemes of relative curves

4.1. Let $f : X \to S$ be a smooth morphism of schemes of relative dimension 1, with connected geometric fibers of genus g, which is Zariski-locally projective over S.

Proposition 4.2. *The canonical homomorphism $\mathcal{O}_S \to f_*\mathcal{O}_X$ is an isomorphism.*

If S is locally noetherian, then $\mathcal{O}_X$ is cohomologically flat over S in dimension 0 by [EGA III_2 1963, 7.8.6]. This means that for any quasicoherent $\mathcal{O}_S$-module $\mathcal{M}$, the canonical homomorphism $f_*f^*\mathcal{O}_X \otimes_{\mathcal{O}_S} \mathcal{M} \to f_*f^*\mathcal{M}$ is an isomorphism. This implies that the formation of $f_*\mathcal{O}_X$ commutes with arbitrary base change: if $f' : X \times_S S' \to S'$ is the base change of f by a morphism of schemes $S' \to S$, then the canonical morphism $f_*\mathcal{O}_X \otimes_{\mathcal{O}_S} \mathcal{O}_{S'} \to f'_*\mathcal{O}_{X\times_S S'}$ is an isomorphism, see [EGA III_2 1963, 7.7.5.3]. By applying this result to the inclusion $\mathrm{Spec}(\kappa(s)) \to S$ of a point s of S, we obtain that $f_*(\mathcal{O}_X)_s \otimes_{\mathcal{O}_{S,s}} \kappa(s)$ is isomorphic to $H^0(X_s, \mathcal{O}_{X_s}) = \kappa(s)$. Since $f_*(\mathcal{O}_X)$ is a coherent $\mathcal{O}_S$-module, Nakayama's lemma yields that the canonical morphism $\mathcal{O}_S \to f_*(\mathcal{O}_X)$ is an epimorphism. It is also injective since f is faithfully flat, hence the result.

In general one can assume that S is affine and that X is projective over S, in which case there is a noetherian scheme S_0, a morphism $S \to S_0$ and a smooth projective S_0-scheme X_0 with geometrically connected fibers such that X is isomorphic to the S-scheme $X_0 \times_{S_0} S$, see [EGA IV$_3$ 1966, 8.9.1, 8.10.5(xiii); EGA IV$_4$ 1967, 17.7.9]. We have already seen that in this case the canonical homomorphism $\mathcal{O}_{S_0} \to f_*\mathcal{O}_{X_0}$ is an isomorphism, and that the formation of $f_*\mathcal{O}_{X_0}$ commutes with arbitrary base change. In particular, both morphisms in the sequence

$$\mathcal{O}_S \to f_*\mathcal{O}_{X_0} \otimes_{\mathcal{O}_{S_0}} \mathcal{O}_S \to f_*\mathcal{O}_X$$

are isomorphisms.

Proposition 4.3. *Let $d \geq 2g-1$ be an integer, and let $\mathcal{L}$ be an invertible $\mathcal{O}_X$-module with degree d on each fiber of f. Then, the $\mathcal{O}_S$-module $f_*\mathcal{L}$ is locally free of rank $d-g+1$, the higher direct images $R^j(f_*\mathcal{L})$ vanish for $j > 0$, and the formation of $f_*\mathcal{L}$ commutes with arbitrary base change: if $f' : X' \to S'$ is the base change of f by a morphism $S' \to S$, then the canonical homomorphism $f_*\mathcal{L} \otimes_{\mathcal{O}_S} \mathcal{O}_{S'} \to f'_*(\mathcal{L} \otimes_{\mathcal{O}_X} \mathcal{O}_{X'})$ is an isomorphism.*

We first assume that S is locally noetherian. For each point of s of S and for each integer i, the Riemann–Roch theorem for smooth projective curves implies that the $k(s)$-vector space $H^i(X_s, \mathcal{L}_s)$ is of dimension $d-g+1$ for $i=0$, and vanishes otherwise. This implies that $R^j f_*(\mathcal{L} \otimes_{\mathcal{O}_X} f^*\mathcal{N})$ vanishes for any integer $j > 0$ and any $\mathcal{O}_S$-module $\mathcal{N}$ by the proof of [EGA III$_2$ 1963, 7.9.8]. Let

$$0 \to \mathcal{N} \to \mathcal{M} \to \mathcal{P} \to 0$$

be an exact sequence of $\mathcal{O}_S$-modules. Since f is flat and since $\mathcal{L}$ is a flat $\mathcal{O}_X$-module, the sequence

$$0 \to \mathcal{L} \otimes_{\mathcal{O}_X} f^*\mathcal{N} \to \mathcal{L} \otimes_{\mathcal{O}_X} f^*\mathcal{M} \to \mathcal{L} \otimes_{\mathcal{O}_X} f^*\mathcal{P} \to 0$$

is exact as well. Since $R^1 f_*(\mathcal{L} \otimes_{\mathcal{O}_X} f^*\mathcal{N})$ vanishes, the sequence

$$0 \to f_*(\mathcal{L} \otimes_{\mathcal{O}_X} f^*\mathcal{N}) \to f_*(\mathcal{L} \otimes_{\mathcal{O}_X} f^*\mathcal{M}) \to f_*(\mathcal{L} \otimes_{\mathcal{O}_X} f^*\mathcal{P}) \to 0$$

is exact. The $\mathcal{O}_X$-module $\mathcal{L}$ is therefore cohomologically flat over S in dimension 0, see [EGA III$_2$ 1963, 7.8.1]. By [EGA III$_2$ 1963, 7.8.4(d)] the $\mathcal{O}_S$-module $f_*\mathcal{L}$ is locally free, and the formation of $f_*\mathcal{L}$ commutes with arbitrary base change. By applying the latter result to the inclusion $\operatorname{Spec}(\kappa(s)) \to S$ of a point s of S and by using that $H^0(X_s, \mathcal{L}_s)$ is of dimension $d-g+1$ over $\kappa(s)$, we obtain that the locally free $\mathcal{O}_X$-module $f_*\mathcal{L}$ is of constant rank $d-g+1$.

In general one can assume that S is affine and that X is projective over S, in which case there is a noetherian scheme S_0, a morphism $S \to S_0$, a smooth projective S_0-scheme X_0, and an invertible $\mathcal{O}_{X_0}$-module $\mathcal{L}_0$ such that X is isomorphic to the S-scheme $X_0 \times_{S_0} S$ and $\mathcal{L}$ is isomorphic to the pullback of $\mathcal{L}_0$ by the canonical projection $X_0 \times_{S_0} S \to X_0$, see [EGA IV$_3$ 1966, 8.9.1, 8.10.5(xiii); EGA IV$_4$ 1967, 17.7.9]. We have seen that the $\mathcal{O}_{S_0}$-module $f_{0*}\mathcal{L}$ is locally free of rank $d-g+1$, and that its formation commutes with arbitrary base change. By performing the base change by the morphism $S \to S_0$, we obtain that $f_*\mathcal{L}$ is a locally free $\mathcal{O}_S$-module of rank $d-g+1$ and that the formation of $f_*\mathcal{L}$ commutes with arbitrary base change.

4.4. Let $f : X \to S$ be as in Section 4.1. The *relative Picard functor* of f is the sheaf of abelian groups $\mathrm{Pic}_S(X) = R^1 f_{\mathrm{Fppf},*}\mathbb{G}_m$ in S_{Fppf}. Alternatively, $\mathrm{Pic}_S(X)$ is the sheaf of abelian groups on S associated to the presheaf which sends an S-scheme T to $\mathrm{Pic}(X \times_S T)$, the abelian group of isomorphism classes of invertible $\mathcal{O}_{X\times_S T}$-modules. For any S-scheme S', we have $(S_{\mathrm{Fppf}})_{/S'} = S'_{\mathrm{Fppf}}$, and we thus have:

Proposition 4.5. *For any S-scheme S', the canonical morphism*

$$\mathrm{Pic}_{S'}(X \times_S S') \to \mathrm{Pic}_S(X) \times_S S'$$

is an isomorphism in S'_{Fppf}.

The elements of $\mathrm{Pic}(X \times_S T)$ which are pulled back from an element of $\mathrm{Pic}(T)$ yield trivial classes in $\mathrm{Pic}_S(X)(T)$, since invertible $\mathcal{O}_T$-modules are locally trivial on T (for the Zariski topology, and thus for the fppf-topology). This yields a sequence

$$0 \to \mathrm{Pic}(T) \to \mathrm{Pic}(X \times_S T) \to \mathrm{Pic}_S(X)(T) \to 0, \tag{4.5.1}$$

which is however not necessarily exact. The following is Proposition 4 from [Bosch et al. 1990, 8.1], whose assumptions are satisfied by Proposition 4.2:

Proposition 4.6. *If f has a section, then the sequence (4.5.1) is exact for any S-scheme T.*

By a theorem of Grothendieck [Bosch et al. 1990, 8.2.1] the sheaf $\mathrm{Pic}_S(X)$ is representable by a separated S-scheme. By [Bosch et al. 1990, 9.3.1] the S-scheme $\mathrm{Pic}_S(X)$ is smooth of relative dimension g, and there is a decomposition

$$\mathrm{Pic}_S(X) = \coprod_{d\in\mathbb{Z}} \mathrm{Pic}^d_S(X),$$

into open and closed subschemes, where $\mathrm{Pic}^d_S(X)$ is the fppf-sheaf associated to the presheaf

$$\begin{aligned} \mathrm{Sch}^{\mathrm{fp}}_{/S} &\to \mathrm{Sets} \\ T &\mapsto \{\mathcal{L} \in \mathrm{Pic}(X \times_S T) \mid \forall \bar{t} \to T, \deg_{X_{\bar{t}}}(\mathcal{L}_{\bar{t}}) = d\}. \end{aligned}$$

Here the condition $\deg_{X_{\bar{t}}}(\mathcal{L}_{\bar{t}}) = d$ runs over all geometric points $\bar{t} \to T$ of T.

4.7. Let $f : X \to S$ be as in Section 4.1, and let $i : Y \hookrightarrow X$ be a closed subscheme of X, which is finite locally free over S of degree $N \geq 1$. A *Y-rigidified line bundle on X* is a pair $(\mathcal{L}, \alpha)$ where $\mathcal{L}$ is a locally free $\mathcal{O}_X$-module of rank 1 and $\alpha : \mathcal{O}_Y \to i^*\mathcal{L}$ is an isomorphism of $\mathcal{O}_Y$-modules. Two Y-rigidified line bundles $(\mathcal{L}, \alpha)$ and $(\mathcal{L}', \alpha')$ are *equivalent* if there is an isomorphism $\beta : \mathcal{L} \to \mathcal{L}'$ of $\mathcal{O}_X$-modules such that $(i^*\beta)\alpha = \alpha'$. If such an isomorphism β exists, then it is unique. Indeed, any other such isomorphism would take the form $\gamma\beta$ for some global section γ of $\mathcal{O}_X^\times$ such that $i^*\gamma = 1$. Since $f_*\mathcal{O}_X = \mathcal{O}_S$ (see Proposition 4.2), we have $\gamma = f^*\delta$ for some global section δ of $\mathcal{O}_S^\times$. Since the restriction of δ along the finite flat surjective morphism $Y \to S$ is trivial, one must have $\delta = 1$ as well, hence $\gamma = 1$.

Proposition 4.8. *Let* $\operatorname{Pic}_S(X, Y)$ *be the presheaf of abelian groups on* $\mathrm{Sch}^{\mathrm{fp}}_{/S}$ *which maps a finitely presented* S*-scheme* T *to the set of isomorphism classes of* Y_T*-rigidified line bundles on* X_T*. Then, the presheaf* $\operatorname{Pic}_S(X, Y)$ *is representable by a smooth separated* S*-scheme of relative dimension* $N + g - 1$.

We first consider the case where $N = 1$:

Lemma 4.9. *The conclusion of Proposition 4.8 holds if* $N = 1$.

Indeed, if $N = 1$ then Y is the image of a section $x : S \to X$ of f. For any finitely presented S-scheme T, we have a morphism

$$\operatorname{Pic}(X \times_S T) \to \operatorname{Pic}_S(X, x)(T), \quad \mathcal{L} \to (\mathcal{L} \otimes (f^*x^*\mathcal{L})^{-1}, \mathrm{id}).$$

The kernel of this homomorphism consists of all invertible $\mathcal{O}_{X\times_S T}$-modules which are given by the pullback of an invertible $\mathcal{O}_T$-module. Moreover, any isomorphism class $(\mathcal{L}, \alpha)$ in $\operatorname{Pic}_S(X, x)(T)$ is the image of $\mathcal{L}$ by this morphism, hence its surjectivity. We conclude by Proposition 4.6 that the canonical projection morphism

$$\operatorname{Pic}_S(X, x) \to \operatorname{Pic}_S(X), \quad (\mathcal{L}, \alpha) \to \mathcal{L},$$

is an isomorphism of presheaves of abelian groups on $\mathrm{Sch}^{\mathrm{fp}}_{/S}$. This yields Lemma 4.9 since $\operatorname{Pic}_S(X)$ is a smooth separated S-scheme of relative dimension g (see Section 4.4).

We now prove Proposition 4.8. Since $X \times_S Y \to Y$ has a section $x = (i \times \mathrm{id}_Y) \circ \Delta_Y$ where $\Delta_Y : Y \to Y \times_S Y$ is the diagonal morphism of Y, we deduce from Lemma 4.9 and its proof that the canonical projection morphism

$$\operatorname{Pic}_Y(X \times_S Y, x) \to \operatorname{Pic}_Y(X \times_S Y) = \operatorname{Pic}_S(X) \times_S Y$$

sending a pair $(\mathcal{L}, \alpha)$ to the class of $\mathcal{L}$ is an isomorphism. Let Z be the Y-scheme $\operatorname{Pic}_Y(X \times_S Y, x)$, and let $(\mathcal{L}_u, \alpha_u)$ be the universal x-rigidified line bundle on $X \times_S Z$. The morphism $Y \times_S Z \to Z$ is finite locally free of rank N, so the pushforward $\mathcal{A}$ of $\mathcal{O}_{Y\times_S Z}$ is a locally free $\mathcal{O}_Z$-algebra of rank N, and the pushforward $\mathcal{M}$ of $i_Z^*\mathcal{L}_u$ is a locally free $\mathcal{O}_Z$-module of rank N. Let $\lambda : \mathcal{M} \to \mathcal{O}_Z$ be the surjective $\mathcal{O}_Z$-linear homomorphism corresponding to $\alpha_u^{-1} : x_Z^*\mathcal{L}_u \to \mathcal{O}_Z$.

Let T be a Y-scheme, and let $(\mathcal{L}, \beta)$ be a Y_T-rigidified line bundle on X_T. The section $x_T : T \to X_T$ uniquely factors through Y_T and we still denote by x_T the corresponding section of Y_T. The pair $(\mathcal{L}, x_T^*\beta)$ is then an x_T-rigidified line bundle on X_T, so that there is a unique morphism $z : T \to Z$ such that $(\mathcal{L}, x_T^*\beta)$ is equivalent to the pullback by z of $(\mathcal{L}_u, \alpha_u)$. Let us assume that $(\mathcal{L}, x_T^*\beta)$ is equal to this pullback. The section β of $i_T^*\mathcal{L}$ over $Y \times_S T$ provides a section $z^*\mathcal{M}$ over T, which we still denote by β, such that $(z^*\lambda)(\beta) = 1$ and $z^*\mathcal{M} = (z^*\mathcal{A})\beta$. Conversely, any such section produces a Y_T-rigidification of $\mathcal{L}$ on X_T. The functor $\mathrm{Pic}_S(X, Y) \times_S Y = \mathrm{Pic}_Y(X \times_S Y, Y \times_S Y)$ is therefore isomorphic to the functor

$$\mathrm{Sch}^{\mathrm{fp}}_{/S} \to \mathrm{Sets}, \quad T \mapsto \{(z, \beta) \mid z \in Z(T), \beta \in \Gamma(T, z^*\mathcal{M}), \lambda(\beta) = 1 \text{ and } \mathcal{M}_T = \mathcal{A}_T\beta\}.$$

This implies that $\mathrm{Pic}_S(X, Y) \times_S Y$ is representable by a relatively affine Z-scheme, smooth of relative dimension $N-1$ over Z. By fppf-descent of affine morphisms of schemes along the fppf-cover $\mathrm{Pic}_S(X) \times_S Y \to \mathrm{Pic}_S(X)$, this implies the representability of $\mathrm{Pic}_S(X, Y)$ by an S-scheme, which is relatively affine and smooth of relative dimension $N-1$ over $\mathrm{Pic}_S(X)$. Since $\mathrm{Pic}_S(X)$ is separated and smooth of relative dimension g over S (see Section 4.1), the S-scheme $\mathrm{Pic}_S(X, Y)$ is separated and smooth of relative dimension $g + N - 1$.

4.10. Let $f : X \to S$ be as in Section 4.1, and let $i : Y \hookrightarrow X$ be a closed subscheme of X, which is finite locally free over S of degree $N \geq 1$. A *Y-trivial effective Cartier divisor of degree d on X* is a pair $(\mathcal{L}, \sigma)$ such that $\mathcal{L}$ is a locally free $\mathcal{O}_X$-module of rank 1 and $\sigma : \mathcal{O}_X \hookrightarrow \mathcal{L}$ is an injective homomorphism such that $i^*\sigma$ is an isomorphism and such that the closed subscheme $V(\sigma)$ of X defined by the vanishing of the ideal $\sigma\mathcal{L}^{-1}$ of $\mathcal{O}_X$ is finite locally free of rank d over S. Two Y-trivial effective divisors $(\mathcal{L}, \sigma)$ and $(\mathcal{L}', \sigma')$ are *equivalent* if there is an isomorphism $\beta : \mathcal{L} \to \mathcal{L}'$ of $\mathcal{O}_X$-modules such that $\beta\sigma = \sigma'$. As in Section 4.7, if such an isomorphism exists then it is unique.

Proposition 4.11. *The map $(\mathcal{L}, \sigma) \mapsto (V(\sigma) \hookrightarrow X)$ is a bijection from the set of equivalence classes of Y-trivial effective Cartiers divisor of degree d on X onto the set of closed subschemes of U which are finite locally free of degree d over S.*

Let $(\mathcal{L}, \sigma)$ be a Y-trivial effective divisor of degree d on X. The ideal $\mathcal{I} = \sigma\mathcal{L}^{-1}$ is an invertible ideal of $\mathcal{O}_X$ such that the vanishing locus $V(\mathcal{I})$ is finite locally free of rank d over S and is contained in U. The pair $(\mathcal{L}, \sigma)$ is equivalent to $(\mathcal{I}^{-1}, 1)$, and $\mathcal{I}$ is uniquely determined by $V(\mathcal{I})$. Conversely for any closed subscheme Z of U which is finite locally free of rank d over S, the scheme Z is proper over S hence closed in X as well, and its defining ideal $\mathcal{I}$ in $\mathcal{O}_{X_T}$ is invertible by [Bosch et al. 1990, 8.2.6(ii)]. The pair $(\mathcal{I}^{-1}, 1)$ is then a Y-trivial effective Cartier divisor of degree d on X.

Proposition 4.12. *Let d be an integer and let $\mathrm{Div}_S^{d,+}(X, Y)$ be the functor which to an S-scheme T associates the set of equivalence classes of Y_T-trivial effective Cartier divisors of degree d on X_T. Then $\mathrm{Div}_S^{d,+}(X, Y)$ is representable by the S-scheme $\mathrm{Sym}_S^d(U)$, the d-th symmetric power of $U = X \setminus Y$ over S* (see Section 2.22). *In particular $\mathrm{Div}_S^{d,+}(X, Y)$ is smooth of relative dimension d over S.*

By Proposition 4.11, the functor $\mathrm{Div}_S^{d,+}(X, Y)$ is isomorphic to the functor which sends an S-scheme T to the set of closed subschemes of U_T which are finite locally free of rank d over T. In other words, $\mathrm{Div}_S^{d,+}(X, Y)$ is isomorphic to the Hilbert functor of d-points in the S-scheme U.

If x is a T-point of U, we denote $\mathcal{O}(-x)$ the kernel of the homomorphism $\mathcal{O}_{X \times_S T} \to x_*\mathcal{O}_T$, which is an invertible ideal sheaf, and by $\mathcal{O}(x)$ its dual, which is endowed with a section $1_x : \mathcal{O}_{X \times_S T} \hookrightarrow \mathcal{O}(x)$.

The morphism

$$\mathrm{Sym}^d_S(U) \to \mathrm{Div}^{d,+}_S(X, Y), \quad (x_1, \dots, x_d) \to \left(\bigotimes_{i=1}^{d} \mathcal{O}(x_i), \prod_{i=1}^{d} 1_{x_i} \right),$$

is then an isomorphism of fppf-sheaves by [SGA 4_3 1973, XVII.6.3.9], hence Proposition 4.12.

Remark 4.13. Let T be an S-scheme. Let Z be a closed subscheme of U_T which is finite locally free of rank d over T, therefore defining a T-point of $\mathrm{Div}^{d,+}_S(X, Y) = \mathrm{Sym}^d_S(U)$ by Proposition 4.11. By [SGA 4_3 1973, XVII.6.3.9], this T-point is given by the composition

$$T \to \mathrm{Sym}^d_T(Z) \to \mathrm{Sym}^d_T(U_T) \to \mathrm{Sym}^d_S(U),$$

where the first morphism is the canonical morphism from Proposition 2.21.

Proposition 4.14. *Let $d \geq N + 2g - 1$ be an integer, and let $\mathrm{Pic}^d_S(X, Y)$ be the inverse image of $\mathrm{Pic}^d_S(X)$ by the natural morphism $\mathrm{Pic}_S(X, Y) \to \mathrm{Pic}_S(X)$. Then the Abel–Jacobi morphism*

$$\Phi_d : \mathrm{Div}^{d,+}_S(X, Y) \to \mathrm{Pic}^d_S(X, Y), \quad (\mathcal{L}, \sigma) \mapsto (\mathcal{L}, i^*\sigma),$$

is surjective smooth of relative dimension $d - N - g + 1$ and it has geometrically connected fibers.

Let Z be the scheme $\mathrm{Pic}^d_S(X, Y)$, and let $(\mathcal{L}_u, \alpha_u)$ be the universal Y-rigidified line bundle of degree d on X_Z. By [Bosch et al. 1990, 8.2.6(ii)], the closed subscheme Y_Z of X_Z is defined by an invertible ideal sheaf $\mathcal{I}$.

Let $\mathcal{E}$ be the pushforward of $\mathcal{M} = \mathcal{L}_u \otimes_{\mathcal{O}_{X_Z}} \mathcal{I}$ by the morphism $f_Z : X_Z \to Z$. By Proposition 4.3, the $\mathcal{O}_Z$-module $\mathcal{E}$ is locally free of rank $d - N - g + 1$, and for any morphism $T \to Z$ the canonical homomorphism

$$\mathcal{E} \otimes_{\mathcal{O}_Z} \mathcal{O}_T \to f_{T*}(\mathcal{M} \otimes_{\mathcal{O}_{X_Z}} \mathcal{O}_{X_T})$$

is an isomorphism, where $f_T : X_T \to T$ is the base change of f by the morphism $T \to S$. We thus obtain an isomorphism

$$E \to E', \tag{4.14.1}$$

of functors on the category of Z-schemes, where E is the functor $T \mapsto \Gamma(T, \mathcal{E} \otimes_{\mathcal{O}_Z} \mathcal{O}_T)$ and E' is the functor $T \mapsto \Gamma(X_T, \mathcal{M} \otimes_{\mathcal{O}_{X_Z}} \mathcal{O}_{X_T})$. Let $\mathcal{F}$ be the pushforward of $\mathcal{L}_u$ by the morphism f_Z. By the same argument, we obtain that the $\mathcal{O}_Z$-module $\mathcal{F}$ is locally free of rank $d - g + 1$, and that we have an isomorphism

$$F \to F', \tag{4.14.2}$$

of functors on the category of Z-schemes, where F is the functor $T \mapsto \Gamma(T, \mathcal{F} \otimes_{\mathcal{O}_Z} \mathcal{O}_T)$ and F' is the functor $T \mapsto \Gamma(X_T, \mathcal{L}_u \otimes_{\mathcal{O}_{X_Z}} \mathcal{O}_{X_T})$. Let us consider the exact sequence

$$0 \to \mathcal{M} \to \mathcal{L}_u \to \mathcal{L}_u \otimes_{\mathcal{O}_{X_Z}} \mathcal{O}_{Y_Z} \to 0.$$

Since $R^1 f_{Z*}\mathcal{M} = 0$ by Proposition 4.3, we obtain an exact sequence

$$0 \to \mathcal{E} \to \mathcal{F} \to \mathcal{G} \to 0,$$

where $\mathcal{G}$ is a locally free $\mathcal{O}_Z$-module of rank N. Together with (4.14.1) and (4.14.2), this yields an exact sequence

$$0 \to E' \to F' \xrightarrow{b} G \to 0,$$

of Z-group schemes in Z_{fppf}, where G is the functor $T \mapsto \Gamma(T, \mathcal{G}_T \otimes_{\mathcal{O}_Z} \mathcal{O}_T)$. The section α_u of $\mathcal{G}$ over Z corresponds to a morphism $\alpha_u : Z \to G$, and we have an isomorphism

$$\mathrm{Div}_S^{d,+}(X, Y) \to F' \times_{b,G,\alpha_u} Z, \quad (\mathcal{L}, \sigma) \mapsto (\sigma, (\mathcal{L}, i^*\sigma)).$$

Since b is an E'-torsor over G in Z_{fppf}, we obtain that $\mathrm{Div}_S^{d,+}(X, Y)$ is an E'-torsor in Z_{fppf}. Since E' is isomorphic to E by (4.14.1), it is smooth of relative dimension $d - N - g + 1$ over Z with geometrically connected fibers, hence the conclusion of Proposition 4.14.

5. Geometric global class field theory

5.1. Let $f : X \to S$ be a smooth morphism of schemes of relative dimension 1, with connected geometric fibers of genus g, which is Zariski-locally projective over S, and let $i : Y \hookrightarrow X$ be a closed subscheme of X which is finite locally free over S of degree $N \geq 1$. Let $j : U \to X$ be the open complement of Y. Let Λ be a finite ring whose cardinality is invertible on S.

Definition 5.2. A locally free Λ-module $\mathcal{F}$ of rank 1 in $U_{\text{Ét}}$ has *ramification bounded by Y over S* if for any geometric point $\bar{x}$ of Y with image $\bar{s}$ in S, the restriction of $\mathcal{F}$ to $\mathrm{Spec}(\widehat{\mathcal{O}_{X_{\bar{s}},\bar{x}}}) \times_{X_{\bar{s}}} U_{\bar{s}}$ has ramification bounded by the multiplicity of $Y_{\bar{s}}$ at $\bar{x}$ (see Definition 3.9).

Theorem 5.3. *Let $\mathcal{F}$ be a locally free Λ-module of rank* 1 *in $U_{\text{Ét}}$ with ramification bounded by Y over S* (see Definition 5.2). *Then, there is a unique* (*up to isomorphism*) *multiplicative locally free Λ-module $\mathcal{G}$ of rank* 1 *on the S-group scheme* $\mathrm{Pic}_S(X, Y)$ (see Remark 2.6) *such that the pullback of $\mathcal{G}$ by the Abel–Jacobi morphism*

$$U \to \mathrm{Pic}_S(X, Y),$$

which sends x to $(\mathcal{O}(x), 1)$, *is isomorphic to $\mathcal{F}$.*

In Section 5.4, we study the restriction of the locally free Λ-module $\mathcal{F}^{[d]}$ of rank 1 on $\mathrm{Div}_S^{d,+}(X, Y)$ (see Proposition 2.32 and Proposition 4.12) to a geometric fiber of the Abel–Jacobi morphism (see Proposition 4.14)

$$\Phi_d : \mathrm{Div}_S^{d,+}(X, Y) \to \mathrm{Pic}_S^d(X, Y), \quad (\mathcal{L}, \sigma) \mapsto (\mathcal{L}, i^*\sigma).$$

This study will enable us to prove Theorem 5.3 in Section 5.10.

5.4. Let k be an algebraically closed field, let X be a smooth connected projective curve of genus g over k and let $i : Y \to X$ be an effective Cartier divisor of degree N with complement U in X. Let $\mathcal{L}$ be a line bundle of degree $d \geq N + 2g - 1$ on X, and let V be the $(d - N - g + 1)$-dimensional affine space over k associated to the k-vector space $\mathcal{V} = H^0(X, \mathcal{L}(-Y))$, i.e., V is the spectrum of the symmetric algebra of the k-module $\mathrm{Hom}_k(\mathcal{V}, k)$. Let τ be a global section of $\mathcal{L}$ on X such that $i^*\tau : \mathcal{O}_Y \to i^*\mathcal{L}$ is an isomorphism.

Proposition 5.5. *Let Λ be a finite ring of cardinality invertible in k, and let $\mathcal{F}$ be a locally free Λ-module of rank 1 in $U_{\text{Ét}}$, with ramification bounded by Y* (see Definition 5.2). *Then the pullback of $\mathcal{F}^{[d]}$* (see Proposition 2.32) *by the morphism*

$$V \to \mathrm{Div}_k^{d,+}(X, Y),$$

which sends a section s of V to $(\mathcal{L}, \tau - s)$, is a constant étale sheaf.

The morphism

$$V \to \mathrm{Div}_k^{d,+}(X, Y),$$

which sends a point σ of V to $(\mathcal{L}, \tau - \sigma)$, is an isomorphism from V to the fiber of Φ_d over the k-point $(\mathcal{L}, i^*\tau)$, see Proposition 4.14. Proposition 5.5 thus implies:

Corollary 5.6. *Let $\mathcal{F}$ be as in Proposition 5.5. Then the locally free Λ-module $\mathcal{F}^{[d]}$ on $\mathrm{Div}_k^{d,+}(X, Y)_{\text{Ét}}$ is constant on the fiber at $(\mathcal{L}, i^*\tau)$ of the morphism*

$$\Phi_d : \mathrm{Div}_k^{d,+}(X, Y) \to \mathrm{Pic}_k^d(X, Y)$$

from Proposition 4.14 .

We now prove Proposition 5.5. To this end, we consider the morphism

$$\psi : \mathbb{A}_V^1 \to \mathrm{Div}_k^{d,+}(X, Y),$$

which sends a pair (t, σ), where t and σ are points of $\mathbb{A}_k^1$ and V respectively, to the point $(\mathcal{L}, \tau - t\sigma)$ of $\mathrm{Div}_k^{d,+}(X, Y)$. Let $\mathcal{F}$ be as in Proposition 5.5, and let $\mathcal{G}$ be the pullback by ψ of $\mathcal{F}^{[d]}$ (see Proposition 2.32). Denoting by $\iota_t : V \to \mathbb{A}_V^1$ the section corresponding to an element t of $k = \mathbb{A}_k^1(k)$, we must prove that the sheaf $\iota_1^{-1}\mathcal{G}$ is constant. The sheaf $\iota_0^{-1}\mathcal{G}$ is constant, since $\psi\iota_0$ is a constant morphism, hence it is sufficient to prove that $\iota_1^{-1}\mathcal{G}$ and $\iota_0^{-1}\mathcal{G}$ are isomorphic. The latter fact follows from the following lemma:

Lemma 5.7. *The locally free Λ-module $\mathcal{G}$ is the pullback of an étale sheaf on V by the projection $\pi : \mathbb{A}_V^1 \to V$.*

We now prove Lemma 5.7. We start by proving that $\mathcal{G}$ is constant on each geometric fiber of the projection π. Since the formation of ψ and $\mathcal{G}$ is compatible with the base change along any field extension of k, it is sufficient to show that $\mathcal{G}$ is constant on each fiber of the projection $\mathbb{A}_V^1 \to V$ at a k-point σ of V. If $\sigma = 0$, then the restriction of ψ to the fiber of π above σ is constant, hence $\mathcal{G}$ is constant on this fiber.

We now assume that σ is nonzero. Since σ vanishes on the nonempty divisor Y and τ does not, the sections σ and τ are k-linearly independent in $H^0(X, \mathcal{L})$. Let D be the greatest divisor on X such that $D \leq \mathrm{div}(\sigma)$ and $D \leq \mathrm{div}(\tau)$. Since the divisor of τ is contained in U, so is D. We can then write $\sigma = \tilde{\sigma} 1_D$ and $\tau = \tilde{\tau} 1_D$, where 1_D is the canonical section of $\mathcal{O}(D)$ and $\tilde{\sigma}, \tilde{\tau}$ are global sections of $\mathcal{L}(-D)$ on X without common zeroes. Thus $f = [\tilde{\tau} : \tilde{\sigma}]$ is a well defined nonconstant morphism from X to $\mathbb{P}^1_k$. Thus, if W is the closed subscheme of $X \times_k \mathbb{A}^1_k$ defined by the vanishing of $\tau - t\sigma$, where t is the coordinate on $\mathbb{A}^1_k$, then we have

$$W = D \times_k \mathbb{A}^1_k \cup (\mathrm{Graph}(f) \cap X \times_k \mathbb{A}^1_k) \hookrightarrow U \times_k \mathbb{A}^1_k.$$

Moreover, the projection $W \to \mathbb{A}^1_k$ is finite flat of degree d, and the restriction of ψ to the fiber at σ factors as

$$\mathbb{A}^1_k \xrightarrow{\varphi} \mathrm{Sym}^d_{\mathbb{A}^1_k}(W) \to \mathrm{Sym}^d_{\mathbb{A}^1_k}(U \times_k \mathbb{A}^1_k) \to \mathrm{Sym}^d_k(U) \to \mathrm{Div}^{d,+}_k(X, Y),$$

where the first morphism φ is obtained from Proposition 2.21, and the last morphism is the isomorphism from Proposition 4.12. Moreover, the pullback of $\mathcal{F}^{[d]}$ to $\mathrm{Sym}^d_{\mathbb{A}^1_k}(W)$ coincides with $(p_1^{-1}\mathcal{F})^{[d]}$, where $p_1 : W \to U$ is the first projection. In particular, the sheaf $\mathcal{G}$ is isomorphic to $\varphi^{-1}(p_1^{-1}\mathcal{F})^{[d]}$.

Set $K = k((t^{-1}))$ and let $\eta = \mathrm{Spec}(K) \to \mathbb{A}^1_k$ be the corresponding punctured formal neighborhood of ∞. Consider the commutative diagram

$$\begin{array}{ccc} \mathbb{A}^1_k & \xrightarrow{\varphi} & \mathrm{Sym}^d_{\mathbb{A}^1_k}(W) \\ \uparrow & & \uparrow \\ \eta & \longrightarrow & \mathrm{Sym}^d_\eta(W \times_{\mathbb{A}^1_k} \eta). \end{array}$$

We can then write

$$W \times_{\mathbb{A}^1_k} \eta = D \times_k \eta \cup \mathrm{Graph}(f) \times_{\mathbb{P}^1_k} \eta = D \times_k \eta \cup X \times_{f, \mathbb{P}^1_k} \eta.$$

The divisors $D \times_k \eta$ and $X \times_{f,\mathbb{P}^1_k} \eta$ of $X \times_k \eta$ are disjoint, since the former lies over closed points of X, while the latter lies over the generic point of X. We thus have a decomposition

$$W \times_{\mathbb{A}^1_k} \eta = D \times_k \eta \amalg X \times_{f,\mathbb{P}^1_k} \eta = \coprod_i \mathrm{Spec}(L_i)$$

where L_i is either of the form $K[T]/(T^{d_i})$ if $\mathrm{Spec}(L_i)$ is a connected component of $D \times_k \eta$, or a field extension of degree d_i of K if $\mathrm{Spec}(L_i)$ is a connected component of $X \times_{f,\mathbb{P}^1_k} \eta$. In the former case, the restriction of $p_1^{-1}\mathcal{F}$ to $\mathrm{Spec}(L_i)$ is constant, while in the latter case, we have the further information that the restriction of $p_1^{-1}\mathcal{F}$ to $\mathrm{Spec}(L_i)$ has ramification bounded by d_i (see Definition 3.9), since the ramification index of f at a point x above ∞ is greater than or equal to the multiplicity of Y at x, and $\mathcal{F}$ has ramification bounded by Y by assumption. Moreover, we have $\sum_i d_i = d$, and the morphism

$\eta \to \mathrm{Sym}^d_\eta(W \times_{\mathbb{A}^1_k} \eta)$ factors through the canonical morphism

$$\prod_i \mathrm{Sym}^{d_i}_\eta(\mathrm{Spec}(L_i)) \to \mathrm{Sym}^d_\eta(W \times_{\mathbb{A}^1_k} \eta).$$

By Proposition 3.13, we obtain that the restriction of $\mathcal{G}$ to η is tamely ramified. Since the tame fundamental group of $\mathbb{A}^1_k$ is trivial, we conclude that $\mathcal{G}$ is a constant étale Λ-module on the fiber of π at σ. The conclusion of Lemma 5.7 then follows from a descent result, namely Lemma 5.9 below.

Remark 5.8. While the proof of Proposition 3.13, which constitutes the core of the proof of Lemma 5.7 above, uses geometric local class field theory, it should be noticed that its statement does not refer to it. This explains why no form of local-global compatibility is required in the proof of Lemma 5.7.

Lemma 5.9. *Let $g : T' \to T$ be a quasicompact smooth compactifiable morphism of schemes of relative dimension δ with geometrically connected fibers, and let $\mathcal{G}$ be an étale sheaf of Λ-modules on $T'_{\text{ét}}$ which is constant on each geometric fiber of g. Then $\mathcal{G}$ is isomorphic to the pullback by g of an étale sheaf of Λ-modules on $T_{\text{ét}}$.*

By [SGA 4_3 1973, XVIII 3.2.5] the functor $Rg_!$ on the derived category of Λ-modules on T admits the functor $g^! : K \mapsto g^*K(\delta)[2\delta]$ as a right adjoint. Let us apply the functor $\mathcal{H}^0$ to the adjunction morphism $\mathcal{G} \to g^! Rg_! \mathcal{G}$. The morphism

$$\mathcal{G} \to \mathcal{H}^0(g^! Rg_! \mathcal{G}) = g^* R^{2\delta} g_! \mathcal{G}(\delta)$$

is an isomorphism, as can be seen by checking the stalks at geometric points with the proper base change theorem.

5.10. We now prove Theorem 5.3. Let $\mathcal{F}$ be a locally free Λ-module of rank 1 over $U_{\text{Ét}}$. The family $(F^{[d]})_{d\geq 0}$ of locally free Λ-modules of rank 1 yields a multiplicative étale Λ-module of rank 1 over the S-semigroup scheme

$$\mathrm{Div}^+_S(X, Y) = \coprod_{d\geq 0} \mathrm{Div}^{d,+}_S(X, Y).$$

For each integer $d \geq N + 2g$, Corollary 5.6 implies that the locally free Λ-module $\mathcal{F}^{[d]}$ of rank 1 on $\mathrm{Div}^{d,+}_S(X, Y)$ (see Propositions Proposition 2.32 and 4.12) is constant on the geometric fibers of the smooth surjective morphism (see Proposition 4.14)

$$\Phi_d : \mathrm{Div}^{d,+}_S(X, Y) \to \mathrm{Pic}^d_S(X, Y), \quad (\mathcal{L}, \sigma) \mapsto (\mathcal{L}, i^*\sigma).$$

This morphism satisfies the conditions of Lemma 5.9 by Proposition 4.14. We can therefore apply Lemma 5.9, and we obtain a locally free Λ-module $\mathcal{G}_d$ of rank 1 over $\mathrm{Pic}^d_S(X, Y)$ such that $\Phi_d^{-1}\mathcal{G}_d$ is isomorphic to $\mathcal{F}^{[d]}$. By Proposition 2.8, the family $(\mathcal{G}_d)_{d\geq N+2g}$ yields a multiplicative locally free Λ-module of rank 1 on the S-semigroup scheme

$$M = \coprod_{d\geq N+2g} \mathrm{Pic}^d_S(X, Y).$$

Since the morphism

$$\rho : M \times_S M \to \mathrm{Pic}_S(X, Y), \quad (x, y) \mapsto xy^{-1},$$

is faithfully flat and quasicompact, we can apply Proposition 2.15, which yields a multiplicative locally free Λ-module $\mathcal{G}$ of rank 1 over $\mathrm{Pic}_S(X, Y)$ whose restriction to $\mathrm{Pic}^d_S(X, Y)$ coincides with $\mathcal{G}_d$ for $d \geq N + 2g$. The families $(\mathcal{F}^{[d]})_{d\geq 0}$ and $(\Phi_d^{-1}\mathcal{G}_d)_{d\geq 0}$ yield multiplicative locally free Λ-modules of rank 1 on the S-semigroup scheme $\mathrm{Div}^+_S(X, Y) = \coprod_{d\geq 0} \mathrm{Div}^{d,+}_S(X, Y)$, whose restrictions to the ideal

$$I = \coprod_{d\geq N+2g} \mathrm{Div}^{d,+}_S(X, Y)$$

of $\mathrm{Div}^+_S(X, Y)$ are isomorphic. We obtain by Proposition 2.7 an isomorphism from $\mathcal{F}^{[d]}$ to $\Phi_d^{-1}\mathcal{G}_d$ for each $d \geq 0$. In particular, the locally free Λ-module $\Phi_1^{-1}\mathcal{G}_1$ of rank 1 is isomorphic to $\mathcal{F}$.

Acknowledgements

This work is part of the author's PhD dissertation and it was prepared at the Institut des Hautes Études Scientifiques and the École Normale Supérieure while the author benefited from their hospitality and support. The author is indebted to Ahmed Abbes for his numerous suggestions and remarks. Further thanks go to Salim Tayou and Christophe Soulé for their comments on this text. The author would also like to thank the anonymous referee for remarks and suggestions which helped to improve this text.

References

[Bosch et al. 1990] S. Bosch, W. Lütkebohmert, and M. Raynaud, *Néron models*, Ergebnisse der Mathematik (3) **21**, Springer, 1990. MR Zbl

[Contou-Carrère 2013] C. Contou-Carrère, "Jacobienne locale d'une courbe formelle relative", *Rend. Semin. Mat. Univ. Padova* **130** (2013), 1–106. MR Zbl

[Deligne 2001] P. Deligne, "Letter to J.-P. Serre, February 8, 1974", 2001. In French. Appendix to S. Bloch and H. Esnault, "Gauss–Manin determinants for rank 1 irregular connections on curves", *Math. Ann* **321**:1 (2001), 15–87. MR Zbl

[EGA 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 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

[Lang 1956] S. Lang, "Sur les séries L d'une variété algébrique", *Bull. Soc. Math. France* **84** (1956), 385–407. MR Zbl

[Laumon 1990] G. Laumon, "Faisceaux automorphes liés aux séries d'Eisenstein", pp. 227–281 in *Automorphic forms, Shimura varieties, and L-functions, I* (Ann Arbor, MI, 1988), edited by L. Clozel and J. S. Milne, Perspect. Math. **10**, Academic Press, Boston, 1990. MR Zbl

[Moret-Bailly 1985] L. Moret-Bailly, *Pinceaux de variétés abéliennes*, Astérisque **129**, Soc. Math. France, Paris, 1985. MR Zbl

[Rosenlicht 1954] M. Rosenlicht, "Generalized Jacobian varieties", *Ann. of Math.* (2) **59** (1954), 505–530. MR Zbl

[Serre 1959] J.-P. Serre, *Groupes algébriques et corps de classes*, Publ. Inst. Math. Univ. Nancago **7**, Hermann, Paris, 1959. MR Zbl

[Serre 1961] J.-P. Serre, "Sur les corps locaux à corps résiduel algébriquement clos", *Bull. Soc. Math. France* **89** (1961), 105–154. MR Zbl

[Serre 1962] J.-P. Serre, *Corps locaux*, Publ. Inst. Math. Univ. Nancago **8**, Hermann, Paris, 1962. MR Zbl

[SGA 1 1971] A. Grothendieck, *Revêtements étales et groupe fondamental* (Séminaire de Géométrie Algébrique du Bois Marie 1960–1961), Lecture Notes in Math. **224**, Springer, 1971. MR Zbl

[SGA 4_3 1973] M. Artin, A. Grothendieck, and J. L. Verdier, *Théorie des topos et cohomologie étale des schémas, Tome 3: Exposés IX–XIX* (Séminaire de Géométrie Algébrique du Bois Marie 1963–1964), Lecture Notes in Math. **305**, Springer, 1973. MR Zbl

[Suzuki 2013] T. Suzuki, "Some remarks on the local class field theory of Serre and Hazewinkel", *Bull. Soc. Math. France* **141**:1 (2013), 1–24. MR Zbl

[Takeuchi 2019] D. Takeuchi, "Blow-ups and class field theory for curves", *Algebra Number Theory* **13**:6 (2019), 1327–1351.

Communicated by Bjorn Poonen
Received 2018-05-21 Revised 2019-02-26 Accepted 2019-04-11

quentin.guignard@ens.fr *Institut des Hautes Etudes Scientifiques, Bures-Sur-Yvette, France*

ALGEBRA AND NUMBER THEORY 13:6 (2019)
dx.doi.org/10.2140/ant.2019.13.1327

Blow-ups and class field theory for curves

Daichi Takeuchi

We propose another proof of geometric class field theory for curves by considering blow-ups of symmetric products of curves.

1. Introduction

Geometric class field theory gives a geometric description of the abelian coverings of a curve by using generalized jacobian varieties. Let us recall its precise statement. Let C be a projective smooth curve over a perfect field k. We assume that C is geometrically connected over k. Fix a modulus $\mathfrak{m}$, i.e., an effective Cartier divisor of C and let U be its complement in C. Denote by $\mathrm{Pic}^0_{C,\mathfrak{m}}$ the corresponding generalized jacobian variety. Let $G^0 \to \mathrm{Pic}^0_{C,\mathfrak{m}}$ be an étale isogeny of smooth commutative algebraic groups and $G^1 \to \mathrm{Pic}^1_{C,\mathfrak{m}}$ be a compatible morphism of torsors. We call such a pair $(G^0 \to \mathrm{Pic}^0_{C,\mathfrak{m}}, G^1 \to \mathrm{Pic}^1_{C,\mathfrak{m}})$ a *covering of* $(\mathrm{Pic}^0_{C,\mathfrak{m}}, \mathrm{Pic}^1_{C,\mathfrak{m}})$. A covering $(G^0 \to \mathrm{Pic}^0_{C,\mathfrak{m}}, G^1 \to \mathrm{Pic}^1_{C,\mathfrak{m}})$ is called *connected abelian* if G^0 is connected and $G^0 \to \mathrm{Pic}^0_{C,\mathfrak{m}}$ is an abelian isogeny. There is a natural map from U to $\mathrm{Pic}^1_{C,\mathfrak{m}}$ sending a point of U to its associated invertible sheaf with a trivialization. Geometric class field theory states:

Theorem 1.1. *Let C be a projective smooth geometrically connected curve over a perfect field k. Fix a modulus $\mathfrak{m}$ of C and denote its complement by U. Let $\mathrm{Pic}^0_{C,\mathfrak{m}}$ be the generalized jacobian variety with modulus $\mathfrak{m}$. Then a connected abelian covering $(G^0 \to \mathrm{Pic}^0_{C,\mathfrak{m}}, G^1 \to \mathrm{Pic}^1_{C,\mathfrak{m}})$ pulls back by the natural map $U \to \mathrm{Pic}^1_{C,\mathfrak{m}}$ to a geometrically connected abelian covering of U whose ramification is bounded by $\mathfrak{m}$. Conversely, every such covering is obtained in this way.*

Originally this theorem was proved by M. Rosenlicht [1954]. S. Lang [1956] generalized his results to an arbitrary algebraic variety. Their works are explained in detail in Serre's book [1988].

On the other hand, in 1980s, P. Deligne found another proof for the tamely ramified case by using symmetric powers of curves [Laumon 1990]. The aim of this paper is to complete his proof by considering blow-ups of symmetric powers of curves.

We have learned that Q. Guignard has done similar work [2019].

Actually we prove a variant of Theorem 1.1 now stated.

Theorem 1.2. *There is an isomorphism of groups between the subgroup of $\mathrm{H}^1(U, \mathbb{Q}/\mathbb{Z})$ consisting of a character χ such that $\mathrm{Sw}_P(\chi) \leq n_P - 1$ for all points $P \in \mathfrak{m}$, where n_P is the multiplicity of $\mathfrak{m}$*

MSC2010: primary 11G45; secondary 14H30.
Keywords: geometric class field theory, ramification theory, generalized jacobian.

at P, and the subgroup of $\mathrm{H}^1(\mathrm{Pic}_{C,\mathfrak{m}}, \mathbb{Q}/\mathbb{Z})$ *consisting of* ρ *which is multiplicative, i.e., the self-external product* $\rho \boxtimes 1 + 1 \boxtimes \rho$ *on* $\mathrm{Pic}_{C,\mathfrak{m}} \times_k \mathrm{Pic}_{C,\mathfrak{m}}$ *equals to* $m^*\rho$*, the pullback of* ρ *by the multiplication map* $m : \mathrm{Pic}_{C,\mathfrak{m}} \times_k \mathrm{Pic}_{C,\mathfrak{m}} \to \mathrm{Pic}_{C,\mathfrak{m}}$.

The relation between Theorems 1.1 and 1.2 will be explained in Section 4.

When k is algebraically closed, Theorem 1.2 can be stated as follows. Let ρ be a multiplicative element of $\mathrm{H}^1(\mathrm{Pic}_{C,\mathfrak{m}}, \mathbb{Q}/\mathbb{Z})$. Fix a closed point $P \in \mathrm{Pic}^1_{C,\mathfrak{m}}$. The multiplicativity of ρ implies that, for an integer d, the pullback of ρ^d by the multiplication by dP $\mathrm{Pic}^0_{C,\mathfrak{m}} \to \mathrm{Pic}^d_{C,\mathfrak{m}}$ coincides with ρ^0. In this way, Theorem 1.2 can be restated as follows:

Theorem 1.3. *Assume that k is algebraically closed. Then there is an isomorphism of groups between the subgroup of* $\mathrm{H}^1(U, \mathbb{Q}/\mathbb{Z})$ *consisting of a character* χ *such that* $\mathrm{Sw}_P(\chi) \leq n_P - 1$ *for all points* $P \in \mathfrak{m}$ *and the subgroup of* $\mathrm{H}^1(\mathrm{Pic}^0_{C,\mathfrak{m}}, \mathbb{Q}/\mathbb{Z})$ *consisting of a multiplicative element* ρ^0*, i.e., the self-external product* $\rho^0 \boxtimes 1 + 1 \boxtimes \rho^0$ *on* $\mathrm{Pic}^0_{C,\mathfrak{m}} \times_k \mathrm{Pic}^0_{C,\mathfrak{m}}$ *equals to* $m^*\rho^0$*, the pullback of* ρ^0 *by the multiplication map* $m : \mathrm{Pic}^0_{C,\mathfrak{m}} \times_k \mathrm{Pic}^0_{C,\mathfrak{m}} \to \mathrm{Pic}^0_{C,\mathfrak{m}}$.

Here we summarize the construction of this paper. In Section 2, we recall the definition and properties of (refined) Swan conductors, and make a calculation on the Swan conductors of symmetric products of characters. We construct compactifications of the Abel–Jacobi maps $U^{(d)} \to \mathrm{Pic}^d_{C,\mathfrak{m}}$ and study their properties in Section 3. The main result of this section is that the compactifications can be identified with open subschemes of blow-ups of $C^{(d)}$. In Section 4, we finish the proof of Theorems 1.1 and 1.2 by combining the results in the previous sections.

Throughout this paper, we use the following conventions: We identify an effective Cartier divisor with the associated closed subscheme. For an object defined on a scheme S (e.g., an S-scheme, a locally free sheaf, a vector bundle, and so on) and a S-scheme T, we denote its pullback to T by the same letter, unless there may be ambiguity. We denote the category of S-schemes by Sch/S. For a category $\mathcal{C}$, we call a functor $\mathcal{C}^{op} \to (Set)$, from the opposite category of $\mathcal{C}$ to the category of sets (Set), a presheaf on $\mathcal{C}$.

2. Preliminaries

In this section, we recall basic properties of Witt vectors and refined Swan conductors, and calculate the Swan conductors of symmetric products of characters. Fix a prime number p.

Reminder on the refined Swan conductor. Let A be a ring of characteristic p. Let m be an integer ≥ 0. We denote by $\mathrm{W}_{m+1}(A)$ the ring of Witt vectors of length $m+1$ with coefficients in A, and write its elements as $(a_0, a_1, \ldots, a_m)$. Let $\mathcal{O}_A$ be the structure sheaf of rings on the étale topos of $\mathrm{Spec}(A)$.

Let F be the absolute Frobenius map $\mathcal{O}_A \to \mathcal{O}_A$, i.e., sending $x \mapsto x^p$, and denote the ring homomorphism $\mathrm{W}_{m+1}(\mathcal{O}_A) \to \mathrm{W}_{m+1}(\mathcal{O}_A)$ induced from F by the same letter F. The short exact sequence

$$0 \to \mathbb{Z}/p^{m+1}\mathbb{Z} \to \mathrm{W}_{m+1}(\mathcal{O}_A) \xrightarrow{F-1} \mathrm{W}_{m+1}(\mathcal{O}_A) \to 0$$

of étale sheaves on $\mathrm{Spec}(A)$ defines the boundary map

$$\delta_{m+1,A} : \mathrm{W}_{m+1}(A) \to \mathrm{H}^1(\mathrm{Spec}(A), \mathbb{Z}/p^{m+1}\mathbb{Z}).$$

The boundary map is surjective, hence $\mathrm{W}_{m+1}(A)/\operatorname{Im}(F-1) \xrightarrow{\sim} \mathrm{H}^1(\mathrm{Spec}(A), \mathbb{Z}/p^{m+1}\mathbb{Z})$, the map it induces, is an isomorphism. The boundary map $\delta_{m+1,A}$ is natural in A. In other words, for a morphism $f : A \to B$ of rings of characteristic p, the diagram

$$\begin{array}{ccc} \mathrm{W}_{m+1}(A) & \xrightarrow{\delta_{m+1,A}} & \mathrm{H}^1(\mathrm{Spec}(A), \mathbb{Z}/p^{m+1}\mathbb{Z}) \\ \downarrow & & \downarrow \\ \mathrm{W}_{m+1}(B) & \xrightarrow{\delta_{m+1,B}} & \mathrm{H}^1(\mathrm{Spec}(B), \mathbb{Z}/p^{m+1}\mathbb{Z}) \end{array} \tag{2-1}$$

is commutative, where the vertical maps are the canonical ones induced from f.

Let (R, π) be a DVR of equal characteristic p and K be its field of fractions. Let v_R be its normalized valuation. Let m be an integer ≥ 0. We extend the valuation v_R to $\mathrm{W}_{m+1}(K)$ by setting

$$v_R((a_0, \dots, a_m)) := \min_i \{p^{m-i} v_R(a_i)\}.$$

We define an increasing exhaustive filtration on $\mathrm{W}_{m+1}(K)$ by setting, for $n \in \mathbb{Z}$, $\mathrm{fil}_n \mathrm{W}_{m+1}(K)$ to be the subgroup of $\mathrm{W}_{m+1}(K)$ consisting of elements $(a_0, \dots, a_m)$ such that

$$v_R((a_0, \dots, a_m)) \geq -n.$$

Define an increasing exhaustive filtration $\mathrm{fil}_n \mathrm{H}^1(K, \mathbb{Z}/p^{m+1}\mathbb{Z})$ of $\mathrm{H}^1(K, \mathbb{Z}/p^{m+1}\mathbb{Z})$ by the image of $\mathrm{fil}_n \mathrm{W}_{m+1}(K)$ through the boundary map $\delta_{m+1,K}$.

For any $\chi \in \mathrm{H}^1(K, \mathbb{Z}/p^{m+1}\mathbb{Z})$, the Swan conductor of χ, $\mathrm{Sw}_R(\chi)$, is the smallest integer $n \geq 0$ such that $\chi \in \mathrm{fil}_n \mathrm{H}^1(K, \mathbb{Z}/p^{m+1}\mathbb{Z})$[Brylinski 1983; Kato 1989]. When R is henselian and the residue field is perfect, this is the same as the classical Swan conductor [Kato 1989, Proposition (6.8)].

Lemma 2.1. *Let R and K be as above. Take $\chi \in \mathrm{H}^1(K, \mathbb{Z}/p^{m+1}\mathbb{Z})$.*

(1) *The subgroup $\mathrm{fil}_0 \mathrm{H}^1(K, \mathbb{Z}/p^{m+1}\mathbb{Z})$ of $\mathrm{H}^1(K, \mathbb{Z}/p^{m+1}\mathbb{Z})$ coincides with the image of the map $\mathrm{H}^1(\mathrm{Spec}(R), \mathbb{Z}/p^{m+1}\mathbb{Z}) \to \mathrm{H}^1(K, \mathbb{Z}/p^{m+1}\mathbb{Z})$, i.e., the group of unramified characters.*

(2) *Let $\hat{R}$ be the completion of R and $\hat{K}$ be its field of fractions. Denote the restriction of χ to $\hat{K}$ by $\hat{\chi}$. Then, the equality $\mathrm{Sw}_R(\chi) = \mathrm{Sw}_{\hat{R}}(\hat{\chi})$ holds.*

Proof. (1) This follows from the commutative diagram

$$\begin{array}{ccc} \mathrm{W}_{m+1}(R) & \xrightarrow{\delta_{m+1,R}} & \mathrm{H}^1(\mathrm{Spec}(R), \mathbb{Z}/p^{m+1}\mathbb{Z}) \\ \downarrow & & \downarrow \\ \mathrm{W}_{m+1}(K) & \xrightarrow{\delta_{m+1,K}} & \mathrm{H}^1(K, \mathbb{Z}/p^{m+1}\mathbb{Z}) \end{array} \tag{2-2}$$

and the fact that the upper horizontal arrow in (2-2) is surjective.

(2) The commutative diagram

$$\begin{array}{ccc} \mathrm{W}_{m+1}(K) & \xrightarrow{\delta_{m+1,K}} & \mathrm{H}^1(K,\mathbb{Z}/p^{m+1}\mathbb{Z}) \\ \downarrow & & \downarrow \\ \mathrm{W}_{m+1}(\hat{K}) & \xrightarrow{\delta_{m+1,\hat{K}}} & \mathrm{H}^1(\hat{K},\mathbb{Z}/p^{m+1}\mathbb{Z}) \end{array}$$

implies $\mathrm{Sw}_R(\chi) \geq \mathrm{Sw}_{\hat{R}}(\hat{\chi})$. Let $n = \mathrm{Sw}_{\hat{R}}(\hat{\chi})$. Then there exists a Witt vector $\hat{\alpha} \in \mathrm{fil}_n \mathrm{W}_{m+1}(\hat{K})$ mapping to $\hat{\chi}$. Take $\alpha \in \mathrm{fil}_n \mathrm{W}_{m+1}(K)$ whose components are close enough to those of $\hat{\alpha}$ with respect to the valuation of $\hat{K}$, so that every component of $\hat{\alpha} - \alpha$ (here α is regarded as an element of $\mathrm{W}_{m+1}(\hat{K})$) is in $\hat{R}$. Then, $\delta_{m+1,\hat{K}}(\hat{\alpha} - \alpha)$ is an unramified character by (1). Therefore, $\chi - \delta_{m+1,K}(\alpha)$ is unramified. Again by (1), there exists $\beta \in \mathrm{W}_{m+1}(R)$ such that $\chi - \delta_{m+1,K}(\alpha) = \delta_{m+1,K}(\beta)$, hence the assertion. □

Next we recall refined Swan conductors.

Define $\widehat{\Omega}^1_R$ to be the π-adic completion of the absolute differential module Ω^1_R. Let $\widehat{\Omega}^1_K := \widehat{\Omega}^1_R \otimes_R K$. The canonical map $\widehat{\Omega}^1_R \to \widehat{\Omega}^1_K$ is injective and we usually regard $\widehat{\Omega}^1_R$ as an R-submodule of $\widehat{\Omega}^1_K$ via this map. The R-module $\widehat{\Omega}^1_R(\log)$ is the R-submodule of $\widehat{\Omega}^1_K$ generated by $\widehat{\Omega}^1_R$ and $d\log\pi := d\pi/\pi$. From the definition, the following holds:

Lemma 2.2. *Assume that R is obtained from a smooth scheme over a perfect field by localizing at a point of codimension one. Let $b_1, \dots, b_n$ be a lift of a p-basis of the residue field of R to R. Then, $\widehat{\Omega}^1_R(\log)$ is a $\hat{R}$-free module with a basis $db_1, \dots, db_n, d\log\pi$.* □

For $\omega \in \widehat{\Omega}^1_K$, define $v^{\log}_R(\omega)$ as the largest integer n such that $\omega \in \pi^n \widehat{\Omega}^1_R(\log)$ (we formally put $v^{\log}_R(0) := \infty$). There is a homomorphism $F^m d : \mathrm{W}_{m+1}(K) \to \widehat{\Omega}^1_K$ of groups given by

$$F^m d((a_0, \dots, a_m)) := \sum_i a_i^{p^{m-i}-1} da_i.$$

Define an increasing exhaustive filtration on $\widehat{\Omega}^1_K$ by setting

$$\mathrm{fil}_n \widehat{\Omega}^1_K := \{\omega \in \widehat{\Omega}^1_K \mid v^{\log}_R(\omega) \geq -n\}$$

for $n \in \mathbb{Z}$. From the definitions, the homomorphism $F^m d : \mathrm{W}_{m+1}(K) \to \widehat{\Omega}^1_K$ respects their filtrations. In other words, $v_R(\alpha) \leq v^{\log}_R(F^m d\alpha)$ hold for all $\alpha \in \mathrm{W}_{m+1}(K)$.

Proposition 2.3 [Leal 2018, Proposition 2.8]. *Let n be an integer ≥ 0.*

(1) *There is a unique homomorphism*

$$\mathrm{rsw} : \mathrm{fil}_n \mathrm{H}^1(K, \mathbb{Z}/p^{m+1}\mathbb{Z}) \to \mathrm{fil}_n \widehat{\Omega}^1_K / \mathrm{fil}_{\lfloor n/p \rfloor} \widehat{\Omega}^1_K,$$

called the refined Swan conductor, such that the composition

$$\operatorname{fil}_n \mathrm{W}_{m+1}(K) \to \operatorname{fil}_n \mathrm{H}^1(K, \mathbb{Z}/p^{m+1}\mathbb{Z}) \to \operatorname{fil}_n \widehat{\Omega}^1_K / \operatorname{fil}_{\lfloor n/p \rfloor} \widehat{\Omega}^1_K$$

coincides with $F^m d$.

(2) *For* $\lfloor \frac{n}{p} \rfloor \leq i \leq n$, *the induced map*

$$\operatorname{fil}_n \mathrm{H}^1(K, \mathbb{Z}/p^{m+1}\mathbb{Z}) / \operatorname{fil}_i \mathrm{H}^1(K, \mathbb{Z}/p^{m+1}\mathbb{Z}) \to \operatorname{fil}_n \widehat{\Omega}^1_K / \operatorname{fil}_i \widehat{\Omega}^1_K$$

is injective.

At the end of this subsection, we extend the definition of the Swan conductors for characters in $\mathrm{H}^1(K, \mathbb{Q}/\mathbb{Z})$ as follows.

Let m be an integer ≥ 0. We identify the groups $\mathbb{Z}/p^m\mathbb{Z}$ and $\frac{1}{p^m}\mathbb{Z}/\mathbb{Z}$ via the multiplication by $\frac{1}{p^m}$. In this way, we define a filtration on $\mathrm{H}^1\big(K, \frac{1}{p^m}\mathbb{Z}/\mathbb{Z}\big)$ from that of $\mathrm{H}^1(K, \mathbb{Z}/p^m\mathbb{Z})$. The natural inclusion $\frac{1}{p^m}\mathbb{Z}/\mathbb{Z} \to \frac{1}{p^{m+1}}\mathbb{Z}/\mathbb{Z}$ induces an inclusion

$$\mathrm{H}^1\big(K, \tfrac{1}{p^m}\mathbb{Z}/\mathbb{Z}\big) \to \mathrm{H}^1\big(K, \tfrac{1}{p^{m+1}}\mathbb{Z}/\mathbb{Z}\big)$$

of groups.

Lemma 2.4. *Let* m, n *be integers* ≥ 0. *The equality*

$$\operatorname{fil}_n \mathrm{H}^1\big(K, \tfrac{1}{p^m}\mathbb{Z}/\mathbb{Z}\big) = \mathrm{H}^1\big(K, \tfrac{1}{p^m}\mathbb{Z}/\mathbb{Z}\big) \cap \operatorname{fil}_n \mathrm{H}^1\big(K, \tfrac{1}{p^{m+1}}\mathbb{Z}/\mathbb{Z}\big)$$

of subgroups of $\mathrm{H}^1\big(K, \frac{1}{p^{m+1}}\mathbb{Z}/\mathbb{Z}\big)$ *holds.*

Proof. Fix a separable closure K^s of K. Let $V : \mathrm{W}_m(K^s) \to \mathrm{W}_{m+1}(K^s)$ be the Verschiebung, i.e., the map sending $(a_0, \dots, a_{m-1})$ to $(0, a_0, \dots, a_{m-1})$. We have the following commutative diagram

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \frac{1}{p^m}\mathbb{Z}/\mathbb{Z} & \longrightarrow & \mathrm{W}_m(K^s) & \xrightarrow{F-1} & \mathrm{W}_m(K^s) & \longrightarrow & 0 \\ & & \big\downarrow & & \big\downarrow V & & \big\downarrow V & & \\ 0 & \longrightarrow & \frac{1}{p^{m+1}}\mathbb{Z}/\mathbb{Z} & \longrightarrow & \mathrm{W}_{m+1}(K^s) & \xrightarrow{F-1} & \mathrm{W}_{m+1}(K^s) & \longrightarrow & 0, \end{array} \tag{2-3}$$

here we identify $\frac{1}{p^m}\mathbb{Z}/\mathbb{Z}$ and $\mathbb{Z}/p^m\mathbb{Z}$ as mentioned above. Taking cohomology groups, we get a commutative diagram

$$\begin{array}{ccc} \mathrm{W}_m(K) & \xrightarrow{\delta_{m,K}} & \mathrm{H}^1\big(K, \frac{1}{p^m}\mathbb{Z}/\mathbb{Z}\big) \\ \big\downarrow V & & \big\downarrow \\ \mathrm{W}_{m+1}(K) & \xrightarrow{\delta_{m+1,K}} & \mathrm{H}^1\big(K, \frac{1}{p^{m+1}}\mathbb{Z}/\mathbb{Z}\big). \end{array} \tag{2-4}$$

Since the map $V : \mathrm{W}_m(K) \to \mathrm{W}_{m+1}(K)$ respects the filtrations, the inclusion

$$\operatorname{fil}_n \mathrm{H}^1\big(K, \tfrac{1}{p^m}\mathbb{Z}/\mathbb{Z}\big) \subset \mathrm{H}^1\big(K, \tfrac{1}{p^m}\mathbb{Z}/\mathbb{Z}\big) \cap \operatorname{fil}_n \mathrm{H}^1\big(K, \tfrac{1}{p^{m+1}}\mathbb{Z}/\mathbb{Z}\big)$$

holds. To prove the equality, it suffices to show that the morphism

$$\mathrm{Gr}_n \, \mathrm{H}^1\big(K, \tfrac{1}{p^m}\mathbb{Z}/\mathbb{Z}\big) \to \mathrm{Gr}_n \, \mathrm{H}^1\big(K, \tfrac{1}{p^{m+1}}\mathbb{Z}/\mathbb{Z}\big)$$

is injective for $n \geq 1$, where $\mathrm{Gr}_n := \mathrm{fil}_n / \mathrm{fil}_{n-1}$. We have the following commutative diagram

$$\begin{array}{ccc} \mathrm{Gr}_n \, \mathrm{H}^1\big(K, \tfrac{1}{p^m}\mathbb{Z}/\mathbb{Z}\big) & \xrightarrow{\mathrm{rsw}} & \mathrm{Gr}_n \, \widehat{\Omega}^1_K \\ \downarrow & \nearrow_{\mathrm{rsw}} & \\ \mathrm{Gr}_n \, \mathrm{H}^1\big(K, \tfrac{1}{p^{m+1}}\mathbb{Z}/\mathbb{Z}\big). & & \end{array} \tag{2-5}$$

By Proposition 2.3(2), the refined Swan conductors rsw in (2-5) are injective, hence the assertion. □

We define a filtration on $\mathrm{H}^1(K, \mathbb{Q}_p/\mathbb{Z}_p) = \bigcup_m \mathrm{H}^1\big(K, \tfrac{1}{p^m}\mathbb{Z}/\mathbb{Z}\big)$ by

$$\mathrm{fil}_n \, \mathrm{H}^1(K, \mathbb{Q}_p/\mathbb{Z}_p) = \bigcup_m \mathrm{fil}_n \, \mathrm{H}^1\big(K, \tfrac{1}{p^m}\mathbb{Z}/\mathbb{Z}\big).$$

Let $\chi \in \mathrm{H}^1(K, \mathbb{Q}/\mathbb{Z})$ be a character. Let χ_p be the p-primary part of χ and be considered as an element of $\mathrm{H}^1(K, \mathbb{Q}_p/\mathbb{Z}_p)$ via the natural decomposition

$$\mathrm{H}^1(K, \mathbb{Q}/\mathbb{Z}) \cong \bigoplus_q \mathrm{H}^1(K, \mathbb{Q}_q/\mathbb{Z}_q),$$

where q runs through all prime numbers. We define the Swan conductor $\mathrm{Sw}(\chi)$ to be the smallest integer $n \geq 0$ such that $\chi_p \in \mathrm{fil}_n \, \mathrm{H}^1(K, \mathbb{Q}_p/\mathbb{Z}_p)$.

The Swan conductor of a symmetric product. In this subsection, we assume that k is a perfect field of characteristic p.

Let X_1, X_2 be smooth schemes over k. Let Z_1 and Z_2 be smooth irreducible closed subvarieties of X_1 and X_2. Let $\tilde{X}_1$, $\tilde{X}_2$, and $\widetilde{X_1 \times X_2}$ be the blow-ups of X_1, X_2, and $X_1 \times X_2$ along Z_1, Z_2, and $Z_1 \times Z_2$. Denote by R_1, R_2, and R_3 the DVRs at the generic points of the exceptional divisor of $\tilde{X}_1$, $\tilde{X}_2$, and $\widetilde{X_1 \times X_2}$. Let K_i be the field of fractions of R_i for $i = 1, 2, 3$.

Lemma 2.5. (1) *The projections $X_1 \times X_2 \to X_1$ and $X_1 \times X_2 \to X_2$ induce the extensions R_3/R_1 and R_3/R_2 of DVRs, which preserve uniformizers.*

(2) *There is a canonical isomorphism*

$$\widehat{\Omega}^1_{K_3} \cong (\hat{K}_3 \otimes_{\hat{K}_1} \widehat{\Omega}^1_{K_1}) \oplus (\hat{K}_3 \otimes_{\hat{K}_2} \widehat{\Omega}^1_{K_2}).$$

This isomorphism respects the filtrations, i.e., via this isomorphism, $\mathrm{fil}_n \, \widehat{\Omega}^1_{K_3}$ *coincides with*

$$(\hat{R}_3 \otimes_{\hat{R}_1} \mathrm{fil}_n \, \widehat{\Omega}^1_{K_1}) \oplus (\hat{R}_3 \otimes_{\hat{R}_2} \mathrm{fil}_n \, \widehat{\Omega}^1_{K_2}).$$

Proof. Let U be the open subscheme of $\widetilde{X_1 \times X_2}$ obtained by removing the strict transforms of $Z_1 \times X_2$ and $X_1 \times Z_2$. This is the largest open subscheme where the pull-backs of $Z_1 \times X_2$ and $X_1 \times Z_2$ coincide with the exceptional divisor. By the universality of the blow-ups $\tilde{X}_1$ and $\tilde{X}_2$, the projections $U \to X_1$ and $U \to X_2$ induce morphisms $U \to \tilde{X}_1$ and $U \to \tilde{X}_2$, hence a morphism $U \to \tilde{X}_1 \times \tilde{X}_2$ of $X_1 \times X_2$-schemes. Denote by D_1 and D_2 the exceptional divisors of $\tilde{X}_1$ and $\tilde{X}_2$. Let $(\tilde{X}_1 \times \tilde{X}_2)'$ be the blow-up of $\tilde{X}_1 \times \tilde{X}_2$ along $D_1 \times D_2$. The morphism $U \to \tilde{X}_1 \times \tilde{X}_2$ lifts to a morphism $U \to (\tilde{X}_1 \times \tilde{X}_2)'$. We claim that this is an open immersion. Indeed, by the universality of the blow-up, the morphism $(\tilde{X}_1 \times \tilde{X}_2)' \to X_1 \times X_2$ lifts to a morphism $(\tilde{X}_1 \times \tilde{X}_2)' \to \widetilde{X_1 \times X_2}$, which implies that U is quasifinite over $(\tilde{X}_1 \times \tilde{X}_2)'$. By Zariski main theorem, the morphism $U \to (\tilde{X}_1 \times \tilde{X}_2)'$ is an open immersion.

Taking an affine open neighborhood of the generic point of the exceptional divisors D_1 and D_2 in $\tilde{X}_1$ and $\tilde{X}_2$, we may assume that $\tilde{X}_1 = \mathrm{Spec}(A_1)$ and $\tilde{X}_2 = \mathrm{Spec}(A_2)$ are affine. We also assume that there are systems of regular parameters $x_1, x_2, \dots, x_n \in A_1$ and $y_1, y_2, \dots, y_m \in A_2$ such that the ideal generated by x_1 and y_1 define D_1 and D_2. The scheme U is canonically isomorphic to $\mathrm{Spec}\big(A_1 \otimes A_2\big[\frac{x_1}{y_1}, \frac{y_1}{x_1}\big]\big)$ and the natural inclusions $A_1, A_2 \to A_1 \otimes A_2\big[\frac{x_1}{y_1}, \frac{y_1}{x_1}\big]$ define the projections $U \to \tilde{X}_1, \tilde{X}_2$. The first assertion follows from this calculation. The canonical isomorphism

$$\Omega^1_{X_1 \times X_2} \cong \mathrm{pr}^*_{X_1} \Omega^1_{X_1} \oplus \mathrm{pr}^*_{X_2} \Omega^1_{X_2},$$

where pr_{X_1} and pr_{X_2} are the projections to X_1 and X_2, gives an isomorphism

$$\widehat{\Omega}^1_{K_3} \cong (\hat{K}_3 \otimes_{\hat{K}_1} \widehat{\Omega}^1_{K_1}) \oplus (\hat{K}_3 \otimes_{\hat{K}_2} \widehat{\Omega}^1_{K_2}).$$

The differentials $\frac{dx_1}{x_1}, d\big(\frac{y_1}{x_1}\big), dx_2, \dots, dx_n, dy_2, \dots, dy_m$ form a basis of $\hat{R}_3$-module $\widehat{\Omega}^1_{R_3}(\log)$. The second assertion follows from this fact and (1). □

Corollary 2.6. *Let* $\chi_i \in \mathrm{H}^1(K_i, \mathbb{Q}/\mathbb{Z})$ *for* $i = 1, 2$. *Then, the following holds*:

$$\mathrm{Sw}_{R_3}(\chi_1 \boxtimes 1 + 1 \boxtimes \chi_2) = \max\{\mathrm{Sw}_{R_1}(\chi_1), \mathrm{Sw}_{R_2}(\chi_2)\}.$$

Proof. Taking the p-primary parts of χ_1, χ_2, and $\chi_1 \boxtimes 1 + 1 \boxtimes \chi_2$, we reduce to the case when $\chi_i \in \mathrm{H}^1(K_i, \mathbb{Z}/p^{m+1}\mathbb{Z})$.

First we verify that the morphism

$$\mathrm{H}^1(K_1, \mathbb{Z}/p^{m+1}\mathbb{Z}) \oplus \mathrm{H}^1(K_2, \mathbb{Z}/p^{m+1}\mathbb{Z}) \to \mathrm{H}^1(K_3, \mathbb{Z}/p^{m+1}\mathbb{Z}) \tag{2-6}$$

respects the filtrations. Since the extensions R_3/R_1, R_3/R_2 of DVRs preserve uniformizers the morphism $\mathrm{W}_{m+1}(K_1) \oplus \mathrm{W}_{m+1}(K_2) \to \mathrm{W}_{m+1}(K_3)$ respects the filtrations, which implies the assertion.

To show the corollary, it is enough to prove that the morphism induced from (2-6) by taking Gr_n is injective. This follows from the injectivity of refined Swan conductors (Proposition 2.3) and Lemma 2.5. □

Let S be a scheme. For a quasiprojective S-scheme X and a natural number $d \geq 1$, the d-th symmetry group $\mathfrak{S}_d$ acts on $X^d := X \times_S X \times_S \cdots \times_S X$ (d times) via permutation of coordinates. Define a scheme $X^{(d)} := X^d/\mathfrak{S}_d$. $X^{(d)}$ is called the d-th symmetric product of X. It is known that, if X is smooth of relative dimension 1 over S, $X^{(d)}$ is smooth and parametrizes effective Cartier divisors of deg $= d$ on X [SGA 4_3 1973, Exposé XVII, Application 1; Polishchuk 2003, 16]. In particular, the formation of $X^{(d)}$ commutes with base change $S' \to S$.

Let C be a projective smooth geometrically connected curve over k. Let U be a nonempty open subscheme of C.

Let d be an integer ≥ 1. We construct a map $\mathrm{H}^1(U, \mathbb{Q}/\mathbb{Z}) \to \mathrm{H}^1(U^{(d)}, \mathbb{Q}/\mathbb{Z})$ as follows. First fix a finite abelian group G. Let $V \to U$ be a G-torsor. Then V^d is a G^d-torsor of U^d. Let H be the subgroup of G^d consisting of elements $(a_1, \ldots, a_d)$ satisfying $\sum_{1\leq i\leq d} a_i = 0$. Then V^d/H is a G-torsor of U^d. This torsor has a natural action by the d-th symmetry group $\mathfrak{S}_d$ which is equivariant with respect to its action to U^d.

Lemma 2.7. *The morphism*

$$(V^d/H)/\mathfrak{S}_d \to U^{(d)} \tag{2-7}$$

induced from the map $V^d/H \to U^d$, taking the quotients by $\mathfrak{S}_d$, is a G-torsor.

Proof. It is sufficient to show that, for every geometric point $\bar{x}$ of U^d, the stabilizer group $(\mathfrak{S}_d)_{\bar{x}}$ at $\bar{x}$ acts trivially on the fiber $(V^d/H)_{\bar{x}}$ over $\bar{x}$, see [SGA 1 1971, Remarque 5.8].

We may assume that k is algebraically closed and that geometric points considered are k-valued points. Let $\bar{x}$ be a geometric point of U^d. For simplicity, we assume that $\bar{x} = (x_1, \ldots, x_1, x_2, \ldots, x_2, \ldots, x_r, \ldots, x_r)$, where $x_1, \ldots, x_r$ are distinct points and x_i appears d_i times for each i. Then the inertia group $(\mathfrak{S}_d)_{\bar{x}}$ at $\bar{x}$ is isomorphic to $\prod_{1\leq i\leq r} \mathfrak{S}_{d_i}$.

For each i, take a k-valued point e_i of $V \times_U x_i$. From the definition of H, the fiber of V^d/H over $\bar{x}$ can be identified with the set

$$\{(e_1, e_1, \ldots, e_r, ge_r) \mid g \in G\}, \tag{2-8}$$

on which $(\mathfrak{S}_d)_{\bar{x}}$ acts trivially. □

In this way, we construct a G-torsor $(V^d/H)/\mathfrak{S}_d$ on $U^{(d)}$. Since this construction is compatible with a morphism of abelian groups $G \to G'$, we obtain a group homomorphism $\mathrm{H}^1(U, \mathbb{Q}/\mathbb{Z}) \to \mathrm{H}^1(U^{(d)}, \mathbb{Q}/\mathbb{Z})$. We denote by $\chi^{(d)}$ the image of χ via this map. Let K be the field of fractions of U, $K_{(d)}$ be that of $U^{(d)}$, and K_d be that of U^d. Taking U smaller and smaller, we also have a map $\mathrm{H}^1(K, G) \to \mathrm{H}^1(K_{(d)}, G)$ for a finite abelian group G and a map $\mathrm{H}^1(K, \mathbb{Q}/\mathbb{Z}) \to \mathrm{H}^1(K_{(d)}, \mathbb{Q}/\mathbb{Z})$.

We consider a similar construction on the groups of Witt vectors. Denote by pr_i^* the morphism $K \to K_d$ induced by the i-th projection $U^d \to U$. Consider the map $\lambda : \mathrm{W}_{m+1}(K) \to \mathrm{W}_{m+1}(K_d)$ sending a Witt vector α to $\mathrm{pr}_1^* \alpha + \cdots + \mathrm{pr}_d^* \alpha$. Since the extension $K_d/K_{(d)}$, induced by the natural projection $U^d \to U^{(d)}$, is finite Galois with the Galois group $\mathfrak{S}_d$, the $\mathfrak{S}_d$-fixed part of $\mathrm{W}_{m+1}(K_d)$ coincides with $\mathrm{W}_{m+1}(K_{(d)})$ (here $\mathrm{W}_{m+1}(K_{(d)})$ is considered as a subgroup of $\mathrm{W}_{m+1}(K_d)$ via the natural projection $U^d \to U^{(d)}$). Thus the map λ factors through $\mathrm{W}_{m+1}(K_{(d)})$. We also denote the induced map $\mathrm{W}_{m+1}(K) \to \mathrm{W}_{m+1}(K_{(d)})$

by λ. Note that the diagram

$$\begin{array}{ccc} \mathrm{W}_{m+1}(K) & \xrightarrow{\delta_{m+1,K}} & \mathrm{H}^1(K, \mathbb{Z}/p^{m+1}\mathbb{Z}) \\ \downarrow \lambda & & \downarrow \\ \mathrm{W}_{m+1}(K_{(d)}) & \xrightarrow{\delta_{m+1,K_{(d)}}} & \mathrm{H}^1(K_{(d)}, \mathbb{Z}/p^{m+1}\mathbb{Z}) \end{array}$$

is commutative. This follows from the commutativity of pr_i^* and the boundary maps (see the diagram (2-1)), and the injectivity of $\mathrm{H}^1(U^{(d)}, \mathbb{Z}/p^{m+1}\mathbb{Z}) \to \mathrm{H}^1(U^d, \mathbb{Z}/p^{m+1}\mathbb{Z})$ (see Lemma 4.2). Also, the canonical morphism $\widehat{\Omega}^1_{K_{(d)}} \otimes_{K_{(d)}} K_d \to \widehat{\Omega}^1_{K_d}$ is an isomorphism and the $\mathfrak{S}_d$-fixed part of $\widehat{\Omega}^1_{K_d}$ coincides with (the image of) $\widehat{\Omega}^1_{K_{(d)}}$. We define a map $\mu : \widehat{\Omega}^1_K \to \widehat{\Omega}^1_{K_{(d)}}$ similarly to λ. The maps λ and μ commute with $F^m d$.

Let P be a closed point of U. For simplicity, let us assume that the residue field at P is isomorphic to k. Let R be the DVR of C at P, and $R_{(d)}$ be the DVR of $K_{(d)}$ at the generic point of the exceptional divisor of the blow-up of $C^{(d)}$ along the point corresponding to the divisor dP. We define filtrations on $\mathrm{W}_{m+1}(K)$ (resp. $\mathrm{W}_{m+1}(K_{(d)})$) and $\widehat{\Omega}^1_K$ (resp. $\widehat{\Omega}^1_{K_{(d)}}$) by R (resp. $R_{(d)}$) (see (2-1)).

The following theorem, and corollary are key calculations to prove Theorem 1.2 in Section 4.

Theorem 2.8. *Let n be an integer.*

(1) *The homomorphism*

$$\lambda : \mathrm{W}_{m+1}(K) \to \mathrm{W}_{m+1}(K_{(d)})$$

sends $\mathrm{fil}_n \mathrm{W}_{m+1}(K)$ *into* $\mathrm{fil}_{\lfloor n/d \rfloor} \mathrm{W}_{m+1}(K_{(d)})$.

(2) *The homomorphism*

$$\mu : \widehat{\Omega}^1_K \to \widehat{\Omega}^1_{K_{(d)}}$$

sends $\mathrm{fil}_n \widehat{\Omega}^1_K$ *into* $\mathrm{fil}_{\lfloor n/d \rfloor} \widehat{\Omega}^1_{K_{(d)}}$. *Let j be an integer. The induced map*

$$\mathrm{fil}_{(j+1)d-1} \widehat{\Omega}^1_K / \mathrm{fil}_{jd-1} \widehat{\Omega}^1_K \to \mathrm{Gr}_j \widehat{\Omega}^1_{K_{(d)}}$$

is injective, here $\mathrm{Gr}_j := \mathrm{fil}_j / \mathrm{fil}_{j-1}$.

Corollary 2.9. *Let χ be a character in $\mathrm{H}^1(K, \mathbb{Q}/\mathbb{Z})$. The following identity holds:*

$$\mathrm{Sw}_{R_{(d)}}(\chi^{(d)}) = \left\lfloor \frac{\mathrm{Sw}_R(\chi)}{d} \right\rfloor.$$

Proof of Corollary 2.9. Taking the p-primary part of χ and an isomorphism $\frac{1}{p^{m+1}}\mathbb{Z}/\mathbb{Z} \cong \mathbb{Z}/p^{m+1}\mathbb{Z}$, we reduce to the case when $\chi \in \mathrm{H}^1(K, \mathbb{Z}/p^{m+1}\mathbb{Z})$. Take $\alpha \in \mathrm{W}_{m+1}(K)$ such that α maps to χ via the boundary map $\mathrm{W}_{m+1}(K) \to \mathrm{H}^1(K, \mathbb{Z}/p^{m+1}\mathbb{Z})$ and $v_R(\alpha) = -\mathrm{Sw}_R(\chi)$. Since the map $F^m d : \mathrm{W}_{m+1}(K) \to \widehat{\Omega}^1_K$ respects the filtrations, we have $F^m d\alpha \in \mathrm{fil}_{-\mathrm{Sw}_R(\chi)} \widehat{\Omega}^1_K$. On the other hand, by Proposition 2.3(2), we have $F^m d\alpha = \mathrm{rsw}(\chi) \notin \mathrm{fil}_{-1-\mathrm{Sw}_R(\chi)} \widehat{\Omega}^1_K$. By the definition of the filtration on $\widehat{\Omega}^1_K$, the equality $v_R^{\log}(F^m d\alpha) = -\mathrm{Sw}_R(\chi)$ holds.

When $\mathrm{Sw}_R(\chi)=0$, χ is unramified since χ is p-torsion. Thus $\chi^{(d)}$ is unramified too by the construction of $\chi^{(d)}$, which implies the assertion in this case.

Assume $\mathrm{Sw}_R(\chi)>0$. Let $r := \lfloor \mathrm{Sw}_R(\chi)/d \rfloor$. From Theorem 2.8(1), the inequality $v_{R_{(d)}}(\lambda(\alpha)) \geq -r$ holds. Thus $\chi^{(d)}$ is contained in $\mathrm{fil}_r\, \mathrm{H}^1(K_{(d)}, \mathbb{Z}/p^{m+1}\mathbb{Z})$, which implies the inequality $\mathrm{Sw}_{R_{(d)}}(\chi^{(d)}) \leq r$.

We show that the class of $\chi^{(d)}$ in $\mathrm{Gr}_r\, \mathrm{H}^1(K_{(d)}, \mathbb{Z}/p^{m+1}\mathbb{Z})$ is nonzero. Consider the following commutative diagram

$$\begin{array}{ccc} \mathrm{fil}_r\, \mathrm{W}_{m+1}(K_{(d)}) & \longrightarrow & \mathrm{Gr}_r\, \mathrm{H}^1(K_{(d)}, \mathbb{Z}/p^{m+1}\mathbb{Z}) \\ & \searrow^{F^m d} & \downarrow \mathrm{rsw} \\ & & \mathrm{Gr}_r\, \widehat{\Omega}^1_{K_{(d)}}, \end{array} \tag{2-9}$$

which is obtained from Proposition 2.3. It suffices to show that $\mathrm{rsw}(\chi^{(d)})$ is nonzero. From the commutativity of (2-9), $\mathrm{rsw}(\chi^{(d)})$ coincides with the class containing $F^m d\lambda(\alpha) = \mu(F^m d\alpha)$. Since $v_R^{\log}(F^m d\alpha) = -\mathrm{Sw}_R(\chi)$, the class of $F^m d\alpha$ in $\mathrm{fil}_{(r+1)d-1}\, \widehat{\Omega}^1_K / \mathrm{fil}_{rd-1}\, \widehat{\Omega}^1_K$ is nonzero. the assertion follows from Theorem 2.8(2), i.e., the injectivity of μ. □

To prove Theorem 2.8, we first collect some basic properties of the DVR $R_{(d)}$ and its module of differentials. Let R_d be the normalization of $R_{(d)}$ in K_d. R_d is a DVR. The natural projection $C^d \to C^{(d)}$ and the i-th projection $C^d \to C$ define extensions of DVRs

$$R_{(d)} \hookrightarrow R_d \xleftarrow{\mathrm{pr}_i^*} R.$$

Fix a uniformizer t of R. Let $S_1, \dots, S_d$ be the elementary symmetric polynomials of $\mathrm{pr}_1^* t, \dots \mathrm{pr}_d^* t$ in R_d, i.e., $S_1, \dots, S_d$ satisfy the following identity

$$(T - \mathrm{pr}_1^* t) \cdots (T - \mathrm{pr}_d^* t) = T^d - S_1 T^{d-1} + \cdots + (-1)^d S_d.$$

Lemma 2.10. (1) *The residue field of $R_{(d)}$ is isomorphic to $k\big(S_1/S_d, \dots, S_{d-1}/S_d\big)$.*

(2) *The elements $S_1, \dots, S_d$ are uniformizers of $R_{(d)}$.*

(3) *The valuations of $\mathrm{pr}_1^* t, \dots, \mathrm{pr}_d^* t$ with respect to R_d are the same.*

Proof. Since the sequence $S_1, \dots, S_d$ is a regular system of parameters of the regular local ring of $C^{(d)}$ at the k-rational point dP, the exceptional divisor of the blow-up of $C^{(d)}$ is isomorphic to a projective space over k with homogeneous coordinates $S_1, \dots, S_d$.

(1) This follows from the considerations above.

(2) At the generic point of the exceptional divisor, the elements $S_1, \dots, S_d$ generate the same ideal. Since the exceptional divisor is regular, the assertion follows.

(3) The d-th symmetry group $\mathfrak{S}_d$ acts on R_d permuting the $pr_i^* t$, hence the assertion. □

By Lemmas 2.2 and 2.10(1), $\widehat{\Omega}^1_{R_{(d)}}(\log)$ is an $\hat{R}_{(d)}$-free module with a basis $dS_1/S_d, \dots, dS_d/S_d$.

Lemma 2.11. *For each integer i, define*

$$\omega_i := \frac{d(\mathrm{pr}_1^* t)}{\mathrm{pr}_1^* t^i} + \cdots + \frac{d(\mathrm{pr}_d^* t)}{\mathrm{pr}_d^* t^i} \in \widehat{\Omega}^1_{K_d}.$$

Let j be an integer. Then, the differentials $\omega_{jd+1}, \ldots, \omega_{(j+1)d}$ form an $\hat{R}_{(d)}$-basis of the $\hat{R}_{(d)}$-free module $(1/S_d^j)\widehat{\Omega}^1_{R_{(d)}}(\log)$.

Proof. To avoid notational confusion, we change the notation d to n in this proof.

Since the differentials ω_j are $\mathfrak{S}_n$-invariant, they are indeed contained in $\widehat{\Omega}^1_{K_{(n)}}$.

Suppose $j \geq 0$. Define a polynomial $F(T) := (T - \mathrm{pr}_1^* t) \cdots (T - \mathrm{pr}_n^* t)$. The following equalities hold:

$$\begin{aligned} -dS_1 T^{n-1} + \cdots + (-1)^n dS_n = dF = -F \sum_{1 \leq i \leq n} \frac{d\,\mathrm{pr}_i^* t}{T - \mathrm{pr}_i^* t} \\ = F \sum_{1 \leq i \leq n} \frac{1}{\mathrm{pr}_i^* t} \frac{d\,\mathrm{pr}_i^* t}{1 - T/\mathrm{pr}_i^* t} = F \sum_{r \geq 0} \omega_{r+1} T^r. \end{aligned}$$

Comparing the coefficients of T^r, we obtain equalities

$$\begin{aligned} S_n \omega_1 &= \pm dS_n \\ S_n \omega_2 \pm S_{n-1}\omega_1 &= \pm dS_{n-1} \\ &\vdots \\ S_n \omega_{r+1} + (\text{a linear combination of } \omega_r, \ldots, \omega_{r-n}) &= 0 \quad (r \geq n) \\ &\vdots \end{aligned}$$

The assertion follows by induction on r.

For the case when $j < 0$, take F as $(1 - \mathrm{pr}_1^* tT) \cdots (1 - \mathrm{pr}_n^* tT)$ and argue similarly. □

Proof of Theorem 2.8. (1) Let $e_{R_d/R_{(d)}}$ be the ramification index of $R_d/R_{(d)}$. Let $e_{R_d/R}$ be the ramification index of R_d/R induced by pr_i. By Lemma 2.10, $e_{R_d/R}$ is independent of i. From the definition of the filtrations, the map $\mathrm{pr}_i^* : \mathrm{W}_{m+1}(K) \to \mathrm{W}_{m+1}(K_d)$ sends $\mathrm{fil}_n \mathrm{W}_{m+1}(K)$ into $\mathrm{fil}_{ne_{R_d/R}} \mathrm{W}_{m+1}(K_d)$. Since S_d is a uniformizer of $R_{(d)}$ by Lemma 2.10, the equality

$$de_{R_d/R} = e_{R_d/R_{(d)}}$$

holds. This shows the identity

$$\mathrm{fil}_{\lfloor n/d \rfloor} \mathrm{W}_{m+1}(K_{(d)}) = \mathrm{fil}_{ne_{R_d/R}} \mathrm{W}_{m+1}(K_d) \cap \mathrm{W}_{m+1}(K_{(d)}),$$

hence the assertion.

(2) Note that the map $\mu : \widehat{\Omega}^1_K \to \widehat{\Omega}^1_{K_{(d)}}$ is continuous. The differentials $dt/t^{n+1}, dt/t^n, \ldots \in \widehat{\Omega}^1_K$ map to $\omega_{n+1}, \omega_n, \ldots$ via μ, all of which are contained in $\mathrm{fil}_{\lfloor n/d \rfloor} \widehat{\Omega}^1_{K_{(d)}}$ by Lemma 2.11. Thus the map μ sends $\mathrm{fil}_n \widehat{\Omega}^1_K$ into $\mathrm{fil}_{\lfloor n/d \rfloor} \widehat{\Omega}^1_{K_{(d)}}$. Since the classes of $dt/t^{(j+1)d}, \ldots, dt/t^{jd+1}$ form a k-basis of $\mathrm{fil}_{(j+1)d-1} \widehat{\Omega}^1_K / \mathrm{fil}_{jd-1} \widehat{\Omega}^1_K$, the last assertion follows from Lemma 2.11. □

3. Generalized jacobians and blow-ups of symmetric powers

In this section, we fix a base scheme S. Let C be a projective smooth S-scheme whose geometric fibers are connected and of dimension 1. Let $\mathfrak{m}$ be a relative effective Cartier divisor of C/S, i.e., a closed subscheme of C which is finite flat of finite presentation over S. We also call $\mathfrak{m}$ a modulus. Let us denote, for S-schemes T, the projections $C \times_S T \to T$ by the same symbol pr. In this section, we recall and study the notion of generalized jacobian varieties. Let d be an integer and $\mathfrak{m}$ be a modulus. Let T be an S-scheme. Consider a datum $(\mathcal{L}, \psi)$ such that:

- $\mathcal{L}$ is an invertible sheaf of $\deg = d$ on C_T.
- ψ is an isomorphism $\mathcal{O}_{\mathfrak{m}_T} \to \mathcal{L}|_{\mathfrak{m}_T}$.

We say that two such data $(\mathcal{L}, \psi)$ and $(\mathcal{L}', \psi')$ are isomorphic if there exists an isomorphism of invertible sheaves $f : \mathcal{L} \to \mathcal{L}'$ making the following diagram commute:

$$\begin{array}{ccc} \mathcal{O}_{\mathfrak{m}_T} & \xrightarrow{\ \psi\ } & \mathcal{L}|_{\mathfrak{m}_T} \\ & {\scriptstyle \psi'} \searrow \quad \swarrow {\scriptstyle f|_{\mathfrak{m}_T}} & \\ & \mathcal{L}'|_{\mathfrak{m}_T} & \end{array}$$

For an S-scheme T, define a set

$$\operatorname{Pic}^{d,\mathrm{pre}}_{C,\mathfrak{m}}(T) := \{\text{isomorphism classes of } (\mathcal{L}, \psi) \text{ defined as above}\}.$$

$\operatorname{Pic}^{d,\mathrm{pre}}_{C,\mathfrak{m}}$ extends in an obvious way to a presheaf on Sch/S, which we also denote by $\operatorname{Pic}^{d,\mathrm{pre}}_{C,\mathfrak{m}}$. Define $\operatorname{Pic}^{d}_{C,\mathfrak{m}}$ as the étale sheafification of $\operatorname{Pic}^{d,\mathrm{pre}}_{C,\mathfrak{m}}$. Their fundamental properties which we use without proofs are:

(1) $\operatorname{Pic}^{d}_{C,\mathfrak{m}}$ are represented by S-schemes.

(2) When $\mathfrak{m}$ is faithfully flat over S, $\operatorname{Pic}^{d,\mathrm{pre}}_{C,\mathfrak{m}}$ are already étale sheaves.

(3) $\operatorname{Pic}^{0}_{C,\mathfrak{m}}$ is a smooth commutative group S-scheme with geometrically connected fibers.

(4) $\operatorname{Pic}^{d}_{C,\mathfrak{m}}$ are $\operatorname{Pic}^{0}_{C,\mathfrak{m}}$-torsors.

When $\mathfrak{m} = 0$, properties (1) and (3) are proved in [Bosch et al. 1990]. For general $\mathfrak{m}$, they can be proved similarly as in [Bosch et al. 1990, 9.3], or can be deduced from the case when $\mathfrak{m} = 0$ and Lemma 3.1.

$\operatorname{Pic}^{0}_{C,\mathfrak{m}}$ is called the generalized jacobian variety of C with modulus $\mathfrak{m}$. When $\mathfrak{m} = 0$, this is the jacobian variety of C. In this case, we also denote $\operatorname{Pic}^{d}_{C,0}$ by $\operatorname{Pic}^{d}_{C}$. Let $\mathfrak{m}$ and $\tilde{\mathfrak{m}}$ be moduli such that $\mathfrak{m} \subset \tilde{\mathfrak{m}}$. There exists a natural map from $\operatorname{Pic}^{d}_{C,\tilde{\mathfrak{m}}}$ to $\operatorname{Pic}^{d}_{C,\mathfrak{m}}$, restricting ψ. Since $\tilde{\mathfrak{m}}$ is a finite S-scheme, this map is a surjection as a morphism of étale sheaves.

Assume that $C \to S$ has a section. In this case, $\operatorname{Pic}^{d}_{C}$ has an expression as a sheaf as follows [Bosch et al. 1990, 8.1]. Let T be an S-scheme, and $\mathcal{L}_1$ and $\mathcal{L}_2$ are invertible sheaves of $\deg = d$ on C_T. Define an equivalence relation on $\operatorname{Pic}^{d,\mathrm{pre}}_{C}$ such that $\mathcal{L}_1$ and $\mathcal{L}_2$ are equivalent if and only if there exists an invertible sheaf $\mathcal{M}$ on T such that $\mathcal{L}_1 \cong \mathcal{L}_2 \otimes \mathrm{pr}^* \mathcal{M}$. If $C \to S$ has a section, the quotient presheaf of $\operatorname{Pic}^{d,\mathrm{pre}}_{C}$ by this equivalence relation is an étale sheaf and coincides with the étale sheafification of $\mathrm{Pic}_C^{d,\mathrm{pre}}$ via the natural surjection. In particular, the identity map $\mathrm{Pic}_C^d \to \mathrm{Pic}_C^d$ corresponds to an equivalence class of invertible sheaves on $C \times_S \mathrm{Pic}_C^d$. In this paper, we call this class the universal class of invertible sheaves of $\deg = d$.

From now on we fix a modulus $\tilde{\mathfrak{m}}$. We call a modulus $\mathfrak{m}$ a submodulus if $\mathfrak{m} \subset \tilde{\mathfrak{m}}$ holds. Until the last paragraph, we treat the case when submoduli considered are everywhere strictly positive on S. Let $\mathfrak{m}$ be a submodulus which is everywhere strictly positive. Then, $\mathrm{Pic}_{C,\mathfrak{m}}^d$ has an explicit expression as a sheaf, as explained before.

Denote the genus of C by g. This is a locally constant function on S. We consider a condition on an integer d as below:

$$d \geq \max\{2g - 1 + \deg \tilde{\mathfrak{m}}, \deg \tilde{\mathfrak{m}}\}. \tag{3-1}$$

When S is quasicompact, such a d always exists. For an integer d and a submodulus $\mathfrak{m}$, denote $d_{\mathfrak{m}} := d - \deg \tilde{\mathfrak{m}} + \deg \mathfrak{m}$. If d satisfies (3-1), $d_{\mathfrak{m}}$ satisfies (3-1) with $\tilde{\mathfrak{m}}$ replaced by $\mathfrak{m}$.

Fix an integer d satisfying (3-1). Let T be an S-scheme and $\mathcal{L}$ be an invertible sheaf of $\deg = d$ on C_T. For every usual point $t \in T$, $R^1 \mathrm{pr}_*(\mathcal{L}(-\tilde{\mathfrak{m}})|_{C_t})$ and $R^1 \mathrm{pr}_*(\mathcal{L}|_{C_t})$ are zero by Serre duality and a degree argument. In this case, $\mathrm{pr}_* \mathcal{L}(-\tilde{\mathfrak{m}})$ and $\mathrm{pr}_* \mathcal{L}$ are locally free sheaves and their formations commute with any base change, i.e., for any morphism of S-schemes $f : T' \to T$, the base change morphisms $f^* \mathrm{pr}_* \mathcal{L} \to \mathrm{pr}_* f^*\mathcal{L}$ and $f^* \mathrm{pr}_*(\mathcal{L}(-\tilde{\mathfrak{m}})) \to \mathrm{pr}_* f^*(\mathcal{L}(-\tilde{\mathfrak{m}}))$ are isomorphisms. Also $R^1 \mathrm{pr}_* f^*\mathcal{L}$ and $R^1 \mathrm{pr}_* f^*(\mathcal{L}(-\tilde{\mathfrak{m}}))$ are zero.

Let $\mathfrak{m}$ be a submodulus. Denote by U the complement of $\mathfrak{m}$ in C. The Abel–Jacobi map $U^{(d_{\mathfrak{m}})} \to \mathrm{Pic}_{C,\mathfrak{m}}^{d_{\mathfrak{m}}}$ is a map which sends $D \in U^{(d_{\mathfrak{m}})}$ to $(\mathcal{O}_C(D), \iota_D)$, where ι_D is the one induced from the natural identification $\mathcal{O}_{C \setminus D} \cong \mathcal{O}_C(D)|_{C \setminus D}$. In this section, we define and study a compactification $\tilde{C}_{\mathfrak{m}}^{(d_{\mathfrak{m}})}$ of the Abel–Jacobi map by constructing the following commutative diagram of smooth S-schemes:

$$\begin{array}{ccccc}
U^{(d_{\mathfrak{m}})} & \hookrightarrow & \tilde{C}_{\mathfrak{m}}^{(d_{\mathfrak{m}})} & \longrightarrow & \mathrm{Pic}_{C,\mathfrak{m}}^{d_{\mathfrak{m}}} \\
\downarrow & & \downarrow & \square & \downarrow \text{(3-7)} \\
X_{\mathfrak{m}} & \xrightarrow{\cong} & \mathbb{P}(\mathcal{E}_{\mathfrak{m}}) & \longrightarrow & P_{\mathfrak{m}}^{d_{\mathfrak{m}}} \\
& \searrow & \downarrow \text{(3-5)} & & \downarrow \text{(3-2)} \\
& & C^{(d_{\mathfrak{m}})} & & \mathrm{Pic}_C^{d_{\mathfrak{m}}}.
\end{array}$$

The S-scheme $\tilde{C}_{\mathfrak{m}}^{(d_{\mathfrak{m}})}$ has, on the one hand, a clear moduli description, and, on the other hand, can be identified by an open subscheme of a blow-up, which will be denoted by $X_{\mathfrak{m}}$, of $C^{(d_{\mathfrak{m}})}$.

Let $\mathcal{L}$ be an invertible sheaf on C_T for an S-scheme T. Denote $\mathcal{L}/(\mathcal{L}(-\mathfrak{m}))$ by $\mathcal{L}_{\mathfrak{m}}$.

For an S-scheme T, consider a pair $(\mathcal{L}, \phi)$ such that $\mathcal{L}$ is an invertible sheaf of $\deg = d_{\mathfrak{m}}$ on C_T and ϕ is an injection $\mathcal{O}_T \to \mathrm{pr}_* \mathcal{L}_{\mathfrak{m}}$ such that the quotient $\mathrm{pr}_* \mathcal{L}_{\mathfrak{m}}/\mathcal{O}_T$ is locally free. Call such pairs $(\mathcal{L}, \phi)$ and $(\mathcal{L}', \phi')$ isomorphic if there exists an isomorphism $f : \mathcal{L} \xrightarrow{\sim} \mathcal{L}'$ such that the following diagram commutes:

$$\begin{array}{ccc} & \mathcal{O}_T & \\ {\scriptstyle\phi}\swarrow & & \searrow{\scriptstyle\phi'} \\ \mathrm{pr}_* \mathcal{L}_{\mathfrak{m}} & \xrightarrow[\mathrm{pr}_* f]{} & \mathrm{pr}_* \mathcal{L}'_{\mathfrak{m}}. \end{array}$$

Define $P^{d_{\mathfrak{m}}}_{\mathfrak{m}}(T)$ as the set of isomorphism classes of such pairs. This is an étale sheaf on Sch/S. Define a map

$$P^{d_{\mathfrak{m}}}_{\mathfrak{m}} \to \mathrm{Pic}^{d_{\mathfrak{m}}}_C \tag{3-2}$$

by forgetting ϕ. Let X be a scheme, and $\mathcal{F}$ be a locally free sheaf of finite rank on X. We use a contra-Grothendieck notation for a projective space. Thus the X-scheme $\mathbb{P}(\mathcal{F})$ parametrizes invertible subsheaves of $\mathcal{F}$.

Lemma 3.1. *The sheaf $P^{d_{\mathfrak{m}}}_{\mathfrak{m}}$ is represented by a proper smooth S-scheme. Assume that $C \to S$ has a section. Let $\mathcal{L}'$ be a representative invertible sheaf of the universal class. Then, as sheaves on $Sch/\mathrm{Pic}^{d_{\mathfrak{m}}}_C$, $P^{d_{\mathfrak{m}}}_{\mathfrak{m}}$ is isomorphic to the projectivization $\mathbb{P}(\mathrm{pr}_* \mathcal{L}'_{\mathfrak{m}})$ of $\mathrm{pr}_* \mathcal{L}'_{\mathfrak{m}}$.*

Proof. First we consider the case when $C(S)$ is not empty. In this case, $\mathrm{Pic}^{d_{\mathfrak{m}}}_C$ has an explicit expression as a sheaf, as explained before.

Via the map (3-2), we regard $P^{d_{\mathfrak{m}}}_{\mathfrak{m}}$ as a sheaf on $Sch/\mathrm{Pic}^{d_{\mathfrak{m}}}_C$. Fix a representative invertible sheaf $\mathcal{L}'$ of the universal class. Let $\mathcal{N}$ be an element of $\mathbb{P}(\mathrm{pr}_* \mathcal{L}'_{\mathfrak{m}})(T)$, where T is a $\mathrm{Pic}^{d_{\mathfrak{m}}}_C$-scheme. Let $\phi : \mathcal{O}_T \to \mathrm{pr}_*((\mathcal{L}' \otimes \mathrm{pr}^* \mathcal{N}^{-1})_{\mathfrak{m}})$ be a morphism obtained by tensoring the inclusion $\mathcal{N} \hookrightarrow \mathrm{pr}_* \mathcal{L}'_{\mathfrak{m}}$ with $\mathcal{N}^{-1}$. Then, the assignment $\mathcal{N} \mapsto (\mathcal{L}' \otimes \mathrm{pr}^* \mathcal{N}^{-1}, \phi)$ defines a morphism of sheaves on $Sch/\mathrm{Pic}^{d_{\mathfrak{m}}}_C$, $\mathbb{P}(\mathrm{pr}_* \mathcal{L}'_{\mathfrak{m}}) \to P^{d_{\mathfrak{m}}}_{\mathfrak{m}}$. This is an isomorphism. Indeed, we can construct its inverse as follows. Let T be a $\mathrm{Pic}^{d_{\mathfrak{m}}}_C$-scheme and $(\mathcal{L}, \phi)$ be an element of $P^{d_{\mathfrak{m}}}_{\mathfrak{m}}(T)$. Let $a : T \to \mathrm{Pic}^{d_{\mathfrak{m}}}_C$ be the structure map. Then, there exists an invertible sheaf $\mathcal{N}$ on T such that $\mathcal{L} \otimes \mathrm{pr}^* \mathcal{N}$ is isomorphic to $a^* \mathcal{L}'$. Such an $\mathcal{N}$ is unique since $C \to S$ has a section. Then, $\mathcal{N} \xrightarrow{\phi \otimes \mathcal{N}} \mathrm{pr}_*((\mathcal{L} \otimes \mathrm{pr}^* \mathcal{N})_{\mathfrak{m}}) \xrightarrow{\sim} \mathrm{pr}_* a^* \mathcal{L}'_{\mathfrak{m}}$ is an element of $\mathbb{P}(\mathrm{pr}_* \mathcal{L}'_{\mathfrak{m}})(T)$.

Next we consider the general case. As the map $C \to S$ has a section étale locally on S, the sheaf $P^{d_{\mathfrak{m}}}_{\mathfrak{m}}$ is represented, étale locally on S, by a projective space bundle. Since the dual of the canonical line bundle of a projective space bundle is relatively ample, the étale descent is effective. □

Let $(\mathcal{L}, \phi)$ be the universal element on $P^{d_{\mathfrak{m}}}_{\mathfrak{m}}$. Define $\mathcal{E}_{\mathfrak{m}}$ as the $\mathcal{O}_{P^{d_{\mathfrak{m}}}_{\mathfrak{m}}}$- module fitting in the following diagram of sheaves on $P^{d_{\mathfrak{m}}}_{\mathfrak{m}}$:

$$\begin{array}{ccc} \mathrm{pr}_*(\mathcal{L}(-\mathfrak{m})) & \longrightarrow & 0 \\ \downarrow & & \downarrow \\ \mathcal{E}_{\mathfrak{m}} & \longrightarrow & \mathcal{O}_{P^{d_{\mathfrak{m}}}_{\mathfrak{m}}} \\ \downarrow & & \downarrow{\scriptstyle\phi} \\ \mathrm{pr}_* \mathcal{L} & \xrightarrow[p]{} & \mathrm{pr}_* \mathcal{L}_{\mathfrak{m}}, \end{array} \tag{3-3}$$

where the bottom horizontal arrow is the pushforward of the quotient map and each square is cartesian. Since all the right arrows are locally split injections and $p : \mathrm{pr}_* \mathcal{L} \to \mathrm{pr}_* \mathcal{L}_{\mathfrak{m}}$ is a surjection of locally free sheaves, $\mathcal{E}_{\mathfrak{m}}$ is locally free of finite rank and all the left arrows are locally split injections.

Let $\mathbb{P}(\mathcal{E}_{\mathfrak{m}})$ be the projectivization of $\mathcal{E}_{\mathfrak{m}}$. As a sheaf on Sch/S, $\mathbb{P}(\mathcal{E}_{\mathfrak{m}})$ parametrizes triples $(\mathcal{L}, \phi, \mathcal{M})$ such that $(\mathcal{L}, \phi)$ is an element of $P_{\mathfrak{m}}^{d_{\mathfrak{m}}}$ and $\mathcal{M}$ is an invertible subsheaf of $\mathrm{pr}_* \mathcal{L} \oplus \mathcal{O}_T$ such that the following diagram commutes:

$$\begin{array}{ccc} \mathcal{M} & \longrightarrow & \mathcal{O}_T \\ \downarrow & & \downarrow \phi \\ \mathrm{pr}_* \mathcal{L} & \xrightarrow[p]{} & \mathrm{pr}_* \mathcal{L}_{\mathfrak{m}}, \end{array} \tag{3-4}$$

where each arrow from $\mathcal{M}$ is the composition of the inclusion $\mathcal{M} \hookrightarrow \mathrm{pr}_* \mathcal{L} \oplus \mathcal{O}_T$ with the respecitve projection. This is a proper smooth S-scheme.

Lemma 3.2. *The map* $\mathrm{pr}_*(\mathcal{L}(-\mathfrak{m})) \to \mathcal{E}_{\mathfrak{m}}$ *in* (3-3) *induces a closed immersion* $\mathbb{P}(\mathrm{pr}_* \mathcal{L}(-\mathfrak{m})) \hookrightarrow \mathbb{P}(\mathcal{E}_{\mathfrak{m}})$. *The closed subspace* $\mathbb{P}(\mathrm{pr}_* \mathcal{L}(-\mathfrak{m}))$ *is a hyperplane bundle of* $\mathbb{P}(\mathcal{E}_{\mathfrak{m}})$.

Proof. The assertion follows from the exact sequence

$$0 \to \mathrm{pr}_*(\mathcal{L}(-\mathfrak{m})) \to \mathcal{E}_{\mathfrak{m}} \to \mathcal{O}_{P_{\mathfrak{m}}^{d_{\mathfrak{m}}}} \to 0. \qquad \square$$

As a subsheaf of $\mathbb{P}(\mathcal{E}_{\mathfrak{m}})$, $\mathbb{P}(\mathrm{pr}_* \mathcal{L}(-\mathfrak{m}))$ parametrizes triples $(\mathcal{L}, \phi, \mathcal{M})$ such that the first projection $\mathcal{M} \to \mathrm{pr}_* \mathcal{L}$ factors through $\mathrm{pr}_* \mathcal{L}(-\mathfrak{m})$.

Now we define a map $\mathbb{P}(\mathcal{E}_{\mathfrak{m}}) \to C^{(d_{\mathfrak{m}})}$ of S-schemes taking the homothety class of the left vertical arrow in (3-4).

Let T be an S-scheme and $(\mathcal{L}, \phi, \mathcal{M})$ be an element of $\mathbb{P}(\mathcal{E}_{\mathfrak{m}})(T)$. Since the arrow $\mathcal{E}_{\mathfrak{m}} \to \mathrm{pr}_* \mathcal{L}$ in (3-3) is locally a split injection, the first projection $\mathcal{M} \to \mathrm{pr}_* \mathcal{L}$ is injective and the cokernel is locally free. Since these hold after any base change $t \to T$ from the spectrum of a field, the map $\mathrm{pr}^* \mathcal{M}_t \to \mathcal{L}_t$ is injective for a usual point t of T. Thus $\mathcal{O}_{C_T} \to \mathcal{L} \otimes \mathrm{pr}^* \mathcal{M}^{-1}$ defines an effective Cartier divisor. Since $\deg(\mathcal{L}^{-1} \otimes \mathrm{pr}^* \mathcal{M})$ equals to $-d_{\mathfrak{m}}$, $\mathrm{Spec}(\mathcal{O}_{C_T}/(\mathcal{L}^{-1} \otimes \mathrm{pr}^* \mathcal{M}))$ is finite flat of finite presentation of $\deg = d_{\mathfrak{m}}$ over T by the Riemann–Roch formula.

Let $C^{(d_{\mathfrak{m}})}$ be the $d_{\mathfrak{m}}$-th symmetric product of C, which parametrizes effective Cartier divisors of $\deg = d_{\mathfrak{m}}$ on C. Define a map

$$\mathbb{P}(\mathcal{E}_{\mathfrak{m}}) \to C^{(d_{\mathfrak{m}})} \tag{3-5}$$

sending $(\mathcal{L}, \phi, \mathcal{M})$ to $\mathrm{Spec}(\mathcal{O}_{C_T}/(\mathcal{L}^{-1} \otimes \mathrm{pr}^* \mathcal{M})) \subset C_T$.

Let Z_0 be the closed subscheme of $C^{(d_{\mathfrak{m}})}$ defined by the map $C^{(d - \deg \tilde{\mathfrak{m}})} \to C^{(d_{\mathfrak{m}})}$, adding $\mathfrak{m}$. Let $X_{\mathfrak{m}}$ be the blow-up of $C^{(d_{\mathfrak{m}})}$ along Z_0. We now construct an isomorphism $X_{\mathfrak{m}} \to \mathbb{P}(\mathcal{E}_{\mathfrak{m}})$, by which we will identify them.

We define a map

$$h : X_{\mathfrak{m}} \to \mathbb{P}(\mathcal{E}_{\mathfrak{m}}) \tag{3-6}$$

as follows.

Let D be the universal effective Cartier divisor on $C^{(d_\mathfrak{m})}$. Denote $\mathcal{O}_{C\times_S C^{(d_\mathfrak{m})}}(D)$ by $\mathcal{O}(D)$ and $\mathcal{O}(D)\otimes\mathcal{O}_{\mathfrak{m}\times_S C^{(d_\mathfrak{m})}}$ by $\mathcal{O}(D)_\mathfrak{m}$ for short. The composition of the natural maps $\mathcal{O}_{C\times_S C^{(d_\mathfrak{m})}}\to\mathcal{O}(D)\to\mathcal{O}(D)_\mathfrak{m}$ defines a map of locally free sheaves $\mathcal{O}_{C^{(d_\mathfrak{m})}}\to\mathrm{pr}_*(\mathcal{O}(D)_\mathfrak{m})$ on $C^{(d_\mathfrak{m})}$. After a base change $T\to C^{(d_\mathfrak{m})}$, this map becomes zero if and only if $T\to C^{(d_\mathfrak{m})}$ factors through Z_0. Thus the image of the dual $(\mathrm{pr}_*\,\mathcal{O}(D)_\mathfrak{m})^\vee\to\mathcal{O}_{C^{(d_\mathfrak{m})}}$ of this map is the ideal $\mathcal{I}$ defining Z_0. Let $\mathcal{L}:=\mathcal{O}_{C\times_S X_\mathfrak{m}}(D)\otimes\mathrm{pr}^*(\mathcal{I}\mathcal{O}_{X_\mathfrak{m}})$. Define $\phi:\mathcal{O}_{X_\mathfrak{m}}\to\mathrm{pr}_*(\mathcal{O}_{C\times_S X_\mathfrak{m}}(D)\otimes\mathrm{pr}^*(\mathcal{I}\mathcal{O}_{X_\mathfrak{m}}))_\mathfrak{m}$ to be the morphism obtained from the map $(\mathcal{I}\mathcal{O}_{X_\mathfrak{m}})^{-1}\to\mathrm{pr}_*\,\mathcal{O}_{C\times_S X_\mathfrak{m}}(D)_\mathfrak{m}$ by tensoring $\mathcal{I}\mathcal{O}_{X_\mathfrak{m}}$. Let $\mathcal{I}\mathcal{O}_{X_\mathfrak{m}}\to\mathrm{pr}_*\,\mathcal{L}$ be the map induced from the natural inclusion $\mathcal{O}_{X_\mathfrak{m}}\to\mathrm{pr}_*\,\mathcal{O}_{C\times_S X_\mathfrak{m}}(D)$ by tensoring with $\mathcal{I}\mathcal{O}_{X_\mathfrak{m}}$. This map and the natural inclusion $\mathcal{I}\mathcal{O}_{X_\mathfrak{m}}\to\mathcal{O}_{X_\mathfrak{m}}$ make the sheaf $\mathcal{I}\mathcal{O}_{X_\mathfrak{m}}$ into a subsheaf of $\mathrm{pr}_*\,\mathcal{L}\oplus\mathcal{O}_{X_\mathfrak{m}}$, which makes the diagram (3-4) commutes. The triple $(\mathcal{L},\phi,\mathcal{I}\mathcal{O}_{X_\mathfrak{m}})$ defines a morphism $h:X_\mathfrak{m}\to\mathbb{P}(\mathcal{E}_\mathfrak{m})$. From the construction, h is a morphism over $C^{(d_\mathfrak{m})}$.

Lemma 3.3. (1) *As a subsheaf of $\mathbb{P}(\mathcal{E}_\mathfrak{m})$, $\mathbb{P}(\mathcal{E}_\mathfrak{m})\times_{C^{(d_\mathfrak{m})}}Z_0$ parametrizes triples $(\mathcal{L},\phi,\mathcal{M})$ such that the second projection $\mathcal{M}\to\mathcal{O}$ are zero. As closed subspaces of $\mathbb{P}(\mathcal{E}_\mathfrak{m})$, $\mathbb{P}(\mathcal{E}_\mathfrak{m})\times_{C^{(d_\mathfrak{m})}}Z_0$ and $\mathbb{P}(\mathrm{pr}_*\,\mathcal{L}(-\mathfrak{m}))$ are equal. In particular, $\mathbb{P}(\mathcal{E}_\mathfrak{m})\times_{C^{(d_\mathfrak{m})}}Z_0$ is a smooth divisor of $\mathbb{P}(\mathcal{E}_\mathfrak{m})$.*

(2) *Let V be the complement of Z_0 in $C^{(d_\mathfrak{m})}$. As a subsheaf of $\mathbb{P}(\mathcal{E}_\mathfrak{m})$, $\mathbb{P}(\mathcal{E}_\mathfrak{m})\times_{C^{(d_\mathfrak{m})}}V$ parametrizes triples $(\mathcal{L},\phi,\mathcal{M})$ such that the second projection $\mathcal{M}\to\mathcal{O}$ is an isomorphism. The projection $\mathbb{P}(\mathcal{E}_\mathfrak{m})\times_{C^{(d_\mathfrak{m})}}V\to V$ is an isomorphism and its inverse coincides with the restriction of h to V.*

Proof. We are considering the following diagram:

$$\begin{array}{ccccc}
\mathbb{P}(\mathcal{E}_\mathfrak{m})\times_{C^{(d_\mathfrak{m})}}Z_0 & \longrightarrow & \mathbb{P}(\mathcal{E}_\mathfrak{m}) & \longleftarrow & \mathbb{P}(\mathcal{E}_\mathfrak{m})\times_{C^{(d_\mathfrak{m})}}V \\
\downarrow & & \downarrow & & \downarrow \\
Z_0 & \longrightarrow & C^{(d_\mathfrak{m})} & \longleftarrow & V
\end{array}$$

(1) Let $(\mathcal{L},\phi,\mathcal{M})$ be an element of $\mathbb{P}(\mathcal{E}_\mathfrak{m})(T)$. This maps into Z_0 via the map $\mathbb{P}(\mathcal{E}_\mathfrak{m})\to C^{(d_\mathfrak{m})}$ if and only if the composition of $\mathrm{pr}^*\,\mathcal{M}\to\mathcal{L}\to\mathcal{L}_\mathfrak{m}$ is zero. Since the right vertical arrow of (3-4) is an injection, this occurs if and only if the second projection $\mathcal{M}\to\mathcal{O}_T$ is zero. The second assertion is obvious from the definition and the expression of $\mathbb{P}(\mathrm{pr}_*\,\mathcal{L}(-\mathfrak{m}))$ as a subsheaf. The last assertion is verified for $\mathbb{P}(\mathrm{pr}_*\,\mathcal{L}(-\mathfrak{m}))$ in Lemma 3.2.

(2) Let T be an S-scheme and $(\mathcal{L},\phi,\mathcal{M})$ be an element of $\mathbb{P}(\mathcal{E}_\mathfrak{m})(T)$. Let t be a usual point of T. By (1), the pullback of the projection $\mathcal{M}\to\mathcal{O}_T$ by $t\hookrightarrow T$ is an isomorphism if and only if the image of t by the map

$$T\xrightarrow{(\mathcal{L},\phi,\mathcal{M})}\mathbb{P}(\mathcal{E}_\mathfrak{m})\to C^{(d_\mathfrak{m})}$$

is in V.

Let $p:\mathbb{P}(\mathcal{E}_\mathfrak{m})\times_{C^{(d_\mathfrak{m})}}V\to V$ be the projection. Since $h:X_\mathfrak{m}\to\mathbb{P}(\mathcal{E}_\mathfrak{m})$ is a $C^{(d_\mathfrak{m})}$-morphism, $p\circ h|_V$ is the identity. Let $(\mathcal{L},\phi,\mathcal{M})$ be an element of $\mathbb{P}(\mathcal{E}_\mathfrak{m})\times_{C^{(d_\mathfrak{m})}}V(T)$. Identify $\mathcal{M}$ and $\mathcal{O}_T$ by the second projection. By this rigidification, $(\mathcal{L},\phi,\mathcal{O}_T)$ is determined by the first projection. Thus p is an injection as a morphism of sheaves. The assertion follows. □

After these preparations, we obtain the following:

Theorem 3.4. *The morphism $h : X_{\mathfrak{m}} \to \mathbb{P}(\mathcal{E}_{\mathfrak{m}})$ in (3-6) is an isomorphism.*

Proof. By Lemma 3.3(1) and the universality of blow-ups, there exists a unique map $\mathbb{P}(\mathcal{E}_{\mathfrak{m}}) \to X_{\mathfrak{m}}$ which is a lift of $\mathbb{P}(\mathcal{E}_{\mathfrak{m}}) \to C^{(d_{\mathfrak{m}})}$. Let V be as in Lemma 3.3(2), i.e., the complement of Z_0 in $C^{(d_{\mathfrak{m}})}$. By the lemma, V can be considered as an open subscheme of $\mathbb{P}(\mathcal{E}_{\mathfrak{m}})$. On the other hand, as V and Z_0 are disjoint, V also can be considered as an open subscheme of $X_{\mathfrak{m}}$. Note that the complements of V in $\mathbb{P}(\mathcal{E}_{\mathfrak{m}})$ and $X_{\mathfrak{m}}$ are supports of divisors. Therefore, V is schematically dense in both of $\mathbb{P}(\mathcal{E}_{\mathfrak{m}})$ and $X_{\mathfrak{m}}$. Since the morphisms $X_{\mathfrak{m}} \to \mathbb{P}(\mathcal{E}_{\mathfrak{m}})$ and $\mathbb{P}(\mathcal{E}_{\mathfrak{m}}) \to X_{\mathfrak{m}}$ constructed above induce the identity on V, and both schemes are separated over S, the assertion follows. □

Let T be an S-scheme and $(\mathcal{L}, \psi)$ be an element of $\mathrm{Pic}^{d_{\mathfrak{m}}}_{C,\mathfrak{m}}(T)$. Define ϕ as the composition $\mathcal{O}_T \to \mathrm{pr}_* \mathcal{O}_{\mathfrak{m}_T} \xrightarrow{\mathrm{pr}_* \psi} \mathrm{pr}_* \mathcal{L}_{\mathfrak{m}}$. Then, the assignment $(\mathcal{L}, \psi) \mapsto (\mathcal{L}, \phi)$ defines a morphism

$$\mathrm{Pic}^{d_{\mathfrak{m}}}_{C,\mathfrak{m}} \to P^{d_{\mathfrak{m}}}_{\mathfrak{m}}. \tag{3-7}$$

Lemma 3.5. *The morphism $\mathrm{Pic}^{d_{\mathfrak{m}}}_{C,\mathfrak{m}} \to P^{d_{\mathfrak{m}}}_{\mathfrak{m}}$ in (3-7) is an open immersion. The open subscheme $\mathrm{Pic}^{d_{\mathfrak{m}}}_{C,\mathfrak{m}}$ parametrizes pairs $(\mathcal{L}, \phi)$ such that the maps $\mathcal{O}_C \to \mathcal{L}_{\mathfrak{m}}$ obtained from ϕ by adjunction are surjective.*

Proof. This morphism is an injection of sheaves. Let $(\mathcal{L}, \phi)$ be an element of $P^{d_{\mathfrak{m}}}_{\mathfrak{m}}(T)$. This element is in $\mathrm{Pic}^{d_{\mathfrak{m}}}_{C,\mathfrak{m}}$ if and only if the map $\mathcal{O}_{C_T} \to \mathcal{L}_{\mathfrak{m}}$ obtained from ϕ by adjunction is a surjection. This is an open condition. □

Next, we study behavior of various schemes when one replaces the modulus $\mathfrak{m}$. Let $\mathfrak{m}$ be a submodulus and $\mathfrak{m}' := \tilde{\mathfrak{m}} - \mathfrak{m}$. Define a closed immersion $C^{(d-\deg \mathfrak{m}')} \to C^{(d)}$ by adding $\mathfrak{m}'$. We denote this closed subscheme of $C^{(d)}$ by $Z_{\mathfrak{m}}$. If $\mathfrak{m}_1 \subset \mathfrak{m}_2$, the inclusion $Z_{\mathfrak{m}_1} \subset Z_{\mathfrak{m}_2}$ holds. The closed immersion $Z_{\mathfrak{m}_1} \hookrightarrow Z_{\mathfrak{m}_2}$ is induced by adding $\mathfrak{m}_2 - \mathfrak{m}_1$. This induces a map

$$X_{\mathfrak{m}_1} \hookrightarrow X_{\mathfrak{m}_2} \tag{3-8}$$

of the blow-ups along Z_0. Let $\mathfrak{m}_1$ and $\mathfrak{m}_2$ be submoduli such that $\mathfrak{m}_1 \subset \mathfrak{m}_2$. Define a map $i_{\mathfrak{m}_1,\mathfrak{m}_2} : P^{d_{\mathfrak{m}_1}}_{\mathfrak{m}_1} \to P^{d_{\mathfrak{m}_2}}_{\mathfrak{m}_2}$ by sending $(\mathcal{L}_1, \phi_1)$ to $(\mathcal{L}_1(\mathfrak{m}_2 - \mathfrak{m}_1), \phi)$, where ϕ is the composition of ϕ_1 and the natural injection $\mathrm{pr}_*(\mathcal{L}_1)_{\mathfrak{m}_1} \to \mathrm{pr}_* \mathcal{L}_1(\mathfrak{m}_2 - \mathfrak{m}_1)_{\mathfrak{m}_2}$. The map $i_{\mathfrak{m}_1,\mathfrak{m}_2}$ is a closed immersion.

Proposition 3.6. (1) *Let $\mathfrak{m}_1$ and $\mathfrak{m}_2$ be submoduli such that $\mathfrak{m}_1 \subset \mathfrak{m}_2$. As a subsheaf of $P^{d_{\mathfrak{m}_2}}_{\mathfrak{m}_2}$, $P^{d_{\mathfrak{m}_1}}_{\mathfrak{m}_1}$ parametrizes pairs $(\mathcal{L}_2, \phi_2)$ such that the compositions $\mathcal{O}_C \xrightarrow{\mathrm{pr}^* \phi_2} (\mathcal{L}_2)_{\mathfrak{m}_2} \to (\mathcal{L}_2)_{\mathfrak{m}_2-\mathfrak{m}_1}$ are zero. The commutative diagram*

$$\begin{array}{ccc} X_{\mathfrak{m}_1} & \longrightarrow & P^{d_{\mathfrak{m}_1}}_{\mathfrak{m}_1} \\ \downarrow & & \downarrow \\ X_{\mathfrak{m}_2} & \longrightarrow & P^{d_{\mathfrak{m}_2}}_{\mathfrak{m}_2} \end{array}$$

induced by (3-6), (3-8), and the projections $\mathbb{P}(\mathcal{E}_{\mathfrak{m}_i}) \to P^{d_{\mathfrak{m}_i}}_{\mathfrak{m}_i}$ is a cartesian diagram.

(2) *Assume that a submodulus* $\mathfrak{m}$ *is the sum* $\sum_i \mathfrak{m}_i$ *of submoduli of* $\deg = 1$. *Let* $\mathfrak{m}'_i := \sum_{j\neq i} \mathfrak{m}_j$. *Then, the open subspace* $\mathrm{Pic}^{d_\mathfrak{m}}_{C,\mathfrak{m}}$ *of* $P^{d_\mathfrak{m}}_\mathfrak{m}$ *is the complement of* $P^{d_{\mathfrak{m}'_i}}_{\mathfrak{m}'_i}$ *for all* i.

Proof. (1) The first assertion is obvious from the definition of $i_{\mathfrak{m}_1,\mathfrak{m}_2}$. To prove the second assertion, it is enough to show that $\mathcal{E}_{\mathfrak{m}_1} \cong i^*_{\mathfrak{m}_1,\mathfrak{m}_2}\mathcal{E}_{\mathfrak{m}_2}$ by Theorem 3.4. Let $(\mathcal{L}_i, \phi_i)$ be the universal elements of $P^{d_{\mathfrak{m}_i}}_{\mathfrak{m}_i}$. The pullback of the cartesian diagram

$$\begin{array}{ccc} \mathcal{E}_{\mathfrak{m}_2} & \longrightarrow & \mathcal{O}_{P^{d_{\mathfrak{m}_2}}_{\mathfrak{m}_2}} \\ \downarrow & & \downarrow \\ \mathrm{pr}_* \mathcal{L}_2 & \longrightarrow & \mathrm{pr}_*(\mathcal{L}_2)_{\mathfrak{m}_2} \end{array}$$

by $i_{\mathfrak{m}_1,\mathfrak{m}_2}$ extends to the diagram

$$\begin{array}{ccc} i^*_{\mathfrak{m}_1,\mathfrak{m}_2}\mathcal{E}_{\mathfrak{m}_2} & \longrightarrow & \mathcal{O}_{P^{d_{\mathfrak{m}_1}}_{\mathfrak{m}_1}} \\ \downarrow & & \downarrow \\ \mathrm{pr}_* \mathcal{L}_1 & \longrightarrow & \mathrm{pr}_*(\mathcal{L}_1)_{\mathfrak{m}_1} \\ \downarrow & & \downarrow \\ \mathrm{pr}_*(\mathcal{L}_1(\mathfrak{m}_2 - \mathfrak{m}_1)) & \longrightarrow & \mathrm{pr}_*(\mathcal{L}_1(\mathfrak{m}_2 - \mathfrak{m}_1))_{\mathfrak{m}_2}, \end{array}$$

where the two squares are cartesian diagrams, which shows the assertion.

(2) This follows from Lemma 3.5 and (1). ∎

Define the S-scheme $\tilde{C}^{(d_\mathfrak{m})}_\mathfrak{m}$ as the fibered product

$$\begin{array}{ccc} \tilde{C}^{(d_\mathfrak{m})}_\mathfrak{m} & \longrightarrow & \mathrm{Pic}^{d_\mathfrak{m}}_{C,\mathfrak{m}} \\ \downarrow & & \downarrow \\ X_\mathfrak{m} & \longrightarrow & P^{d_\mathfrak{m}}_\mathfrak{m}, \end{array} \tag{3-9}$$

where the bottom horizontal map $X_\mathfrak{m} \to P^{d_\mathfrak{m}}_\mathfrak{m}$ is the composition $X_\mathfrak{m} \xrightarrow{\sim} \mathbb{P}(\mathcal{E}_\mathfrak{m}) \to P^{d_\mathfrak{m}}_\mathfrak{m}$. The S-scheme $\tilde{C}^{(d_\mathfrak{m})}_\mathfrak{m}$ is a projective space bundle on $\mathrm{Pic}^{d_\mathfrak{m}}_{C,\mathfrak{m}}$.

Proposition 3.7. *The first projection* $\tilde{C}^{(d_\mathfrak{m})}_\mathfrak{m} \to X_\mathfrak{m}$ *is an open immersion. Moreover, if* $\mathfrak{m}$ *is the sum* $\sum_i \mathfrak{m}_i$ *of submoduli of* $\deg = 1$, $\tilde{C}^{(d_\mathfrak{m})}_\mathfrak{m}$ *coincides with the complement of* $X_{\mathfrak{m}'_i}$ *for all* i, *where* $\mathfrak{m}'_i := \sum_{j\neq i} \mathfrak{m}_j$.

Proof. These are consequences of Lemma 3.5 and Proposition 3.6. ∎

The Abel–Jacobi map $U^{(d_\mathfrak{m})} \to \mathrm{Pic}^{d_\mathfrak{m}}_{C,\mathfrak{m}}$ and the canonical open immersion $U^{(d_\mathfrak{m})} \to X_\mathfrak{m}$ define the following commutative diagram

$$\begin{array}{ccc} U^{(d_\mathfrak{m})} & \longrightarrow & \mathrm{Pic}^{d_\mathfrak{m}}_{C,\mathfrak{m}} \\ \downarrow & & \downarrow \\ X_\mathfrak{m} & \longrightarrow & P^{d_\mathfrak{m}}_\mathfrak{m}, \end{array} \tag{3-10}$$

which induces an $X_\mathfrak{m}$-morphism $U^{(d_\mathfrak{m})} \to \tilde{C}^{(d_\mathfrak{m})}_\mathfrak{m}$. This is an open immersion, since the vertical arrows of (3-9) and the left vertical arrow of (3-10) are open immersions. Combining the previous results, we obtain the following:

Corollary 3.8. *As an open subscheme of $\tilde{C}^{(d_\mathfrak{m})}_\mathfrak{m}$, $U^{(d_\mathfrak{m})}$ is the complement of $\tilde{C}^{(d_\mathfrak{m})} \times_{C^{(d_\mathfrak{m})}} Z_0$.*

Proof. After a finite faithfully flat base change of S, we may assume that $\mathfrak{m}$ decomposes the sum $\sum_i \mathfrak{m}_i$ of submoduli $\mathfrak{m}_i$ of $\deg = 1$. Let V be the complement of Z_0 in $C^{(d_\mathfrak{m})}$. Since $V \setminus U^{(d_\mathfrak{m})}$ is included in $\bigcup_i Z_{\mathfrak{m}'_i}$, where $\mathfrak{m}'_i := \sum_{j \neq i} \mathfrak{m}_j$, the assertion follows from Proposition 3.7. □

Now assume $\mathfrak{m} = 0$. If C has an S-valued point P, it is well-known that $C^{(d)}$ is a projective space bundle over Pic^d_C when $d \geq \max\{2g-1, 0\}$, where g is the genus of C. In other words, there exists a locally free sheaf $\mathcal{F}$ of finite rank on Pic^d_C such that $C^{(d)}$ is isomorphic to $\mathbb{P}(\mathcal{F})$. Classically this is proved using the Poincaré bundle. On the other hand, using Proposition 3.7, we might prove this fact with an extra condition $d \geq \max\{2g, 1\}$, identifying $\mathrm{Pic}^d_{C,P} \cong \mathrm{Pic}^d_C$ (see [Bosch et al. 1990, 8.2]).

Corollary 3.9. *Assume that S is connected noetherian. Let $\mathfrak{m}$ be a modulus > 0 (resp. $= 0$) of C and d be a sufficiently large integer. Take a geometric point $\bar{x}$ on $\tilde{C}^{(d)}_\mathfrak{m}$ (resp. on $C^{(d)}$) and denote $\bar{y}$ its image to $\mathrm{Pic}^d_{C,\mathfrak{m}}$. Then, the morphism of profinite groups $\pi_1(\tilde{C}^{(d)}_\mathfrak{m}, \bar{x}) \to \pi_1(\mathrm{Pic}^d_{C,\mathfrak{m}}, \bar{y})$ (resp. $\pi_1(C^{(d)}, \bar{x}) \to \pi_1(\mathrm{Pic}^d_C, \bar{y})$) induced from the projection $\tilde{C}^{(d)}_\mathfrak{m} \to \mathrm{Pic}^d_{C,\mathfrak{m}}$ (resp. $C^{(d)} \to \mathrm{Pic}^d_C$) is an isomorphism.*

Proof. When $\mathfrak{m} = 0$, let us also denote $C^{(d)}$ by $\tilde{C}^{(d)}_\mathfrak{m}$ for. If $\mathfrak{m} > 0$ (resp. $= 0$), $\tilde{C}^{(d)}_\mathfrak{m}$ is a projective space bundle over $\mathrm{Pic}^d_{C,\mathfrak{m}}$ (resp. after the base change from S to an étale cover). In any case, the morphism $\tilde{C}^{(d)}_\mathfrak{m} \to \mathrm{Pic}^d_{C,\mathfrak{m}}$ is proper surjective smooth with geometrically connected fibers. Take a geometric point $\bar{s}$ of $\tilde{C}^{(d)}_{\mathfrak{m},\bar{y}}$ above $\bar{x}$. Since the scheme $\tilde{C}^{(d)}_{\mathfrak{m},\bar{y}}$ is simply connected, the homotopy exact sequence

$$\pi_1(\tilde{C}^{(d)}_{\mathfrak{m},\bar{y}}, \bar{s}) \to \pi_1(\tilde{C}^{(d)}_\mathfrak{m}, \bar{x}) \to \pi_1(\mathrm{Pic}^d_{C,\mathfrak{m}}, \bar{y}) \to 1$$

implies the assertion. □

4. Proofs

In this section, we prove Theorems 1.1 and 1.2.

First we need to recall basic results on symmetric products of curves.

Let C be a projective smooth geometrically connected curve over a perfect field k. Let $\mathfrak{m}$ be a modulus on C and write $\mathfrak{m} = n_1 P_1 + \cdots + n_r P_r$, where $P_1, \ldots, P_r$ are distinct closed points of $\mathfrak{m}$. Denote the complement of $\mathfrak{m}$ in C by U. Let $d_i := \deg P_i$. Take a positive integer d so that $d \geq \deg \mathfrak{m}$.

Lemma 4.1. *The morphism $\pi : C^{(n_1d_1)} \times \cdots \times C^{(n_rd_r)} \times C^{(d-\deg \mathfrak{m})} \to C^{(d)}$, taking the sum, is étale at the generic point of the closed subvariety $\{n_1P_1\} \times \cdots \times \{n_rP_r\} \times C^{(d-\deg \mathfrak{m})}$ of $C^{(n_1d_1)} \times \cdots \times C^{(n_rd_r)} \times C^{(d-\deg \mathfrak{m})}$.*

Proof. We may assume that k is algebraically closed (hence $d_i = 1$ for all i). Since the map $\pi : C^{(n_1)} \times \cdots \times C^{(n_r)} \times C^{(d-\deg \mathfrak{m})} \to C^{(d)}$ is finite flat, it is enough to show that there exists a closed point Q of $n_1P_1 + \cdots n_rP_r + C^{(d-\deg \mathfrak{m})}$ over which there are $\deg \pi$ points on $C^{(n_1)} \times \cdots \times C^{(n_r)} \times C^{(d-\deg \mathfrak{m})}$. Choose Q as a point corresponding to a divisor $n_1P_1 + \cdots n_rP_r + P_{r+1} + \cdots + P_{r+d-\deg \mathfrak{m}}$, where $P_1, \ldots, P_{r+d-\deg \mathfrak{m}}$ are distinct points of $U(k)$. □

Lemma 4.2. *The morphism $\pi_1(U^d) \to \pi_1(U^{(d)})$ induced from the natural projection $U^d \to U^{(d)}$ (base points are omitted) is surjective.*

Proof. Since U^d and $U^{(d)}$ are geometrically connected over k, it is enough to show the surjectivity after the base change to an algebraic closure $\bar{k}$ by considering the homotopy exact sequence $1 \to \pi_1(U^d_{\bar{k}}) \to \pi_1(U^d) \to \pi_1(\mathrm{Spec}(k)) \to 1$ and the counterpart of $U^{(d)}$.

Assume that k is algebraically closed. Let V be a connected finite étale covering of $U^{(d)}$. We show that the pullback $V \times_{U^{(d)}} U^d$ is also connected, which shows the assertion. Note that, since the schemes U^d and $U^{(d)}$ are normal, V and $V \times_{U^{(d)}} U^d$ are normal. In particular, $V \times_{U^{(d)}} U^d$ is the disjoint union of integral schemes. Since the map $U^d \to U^{(d)}$ is finite flat, each connected component of $V \times_{U^{(d)}} U^d$ surjects onto V. Take a k-valued point $P \in U(k)$. Take a k-valued point $P' \in V(k)$ over $dP \in U^{(d)}(k)$. Since the fiber of $U^d \to U^{(d)}$ over the point dP consists of one point $(P, P, \ldots, P)$, the fiber of $V \times_{U^{(d)}} U^d \to V$ over P' also consists of one point. Thus the scheme $V \times_{U^{(d)}} U^d$ has only one connected component. □

Proof of Theorem 1.2. Let C be a projective smooth geometrically connected curve over a perfect field k. Let $\mathfrak{m} = n_1P_1 + \cdots + n_rP_r$ ($n_i \geq 1$) be a modulus on C, and U be its complement. Set A as the subgroup of $\mathrm{H}^1(U, \mathbb{Q}/\mathbb{Z})$ consisting of characters χ such that $\mathrm{Sw}_{P_i}(\chi) \leq n_i - 1$ for $i = 1, \ldots, r$, and B as the subgroup of $\mathrm{H}^1(\mathrm{Pic}_{C,\mathfrak{m}}, \mathbb{Q}/\mathbb{Z})$ consisting of multiplicative elements.

We construct a map $\Psi : B \to A$. Take $\rho \in B$. Define χ to be the pullback of ρ^1 by the natural map $U \to \mathrm{Pic}^1_{C,\mathfrak{m}}$. We need to show that the ramification is bounded by $\mathfrak{m}$. Take a natural number d large enough so that d satisfies (3-1) for $\mathfrak{m}$. Consider the following commutative diagram:

$$\begin{array}{ccc} U^d = U \times \cdots \times U & \xrightarrow{\pi} & U^{(d)} \\ \downarrow & & \downarrow p \\ \mathrm{Pic}^1_{C,\mathfrak{m}} \times \cdots \times \mathrm{Pic}^1_{C,\mathfrak{m}} & \longrightarrow & \mathrm{Pic}^d_{C,\mathfrak{m}} \end{array} \tag{4-1}$$

By the multiplicativity of ρ, we know that $\pi^*p^*\rho^d = \chi^{\boxtimes d}$. Lemma 4.2 implies that $p^*\rho^d = \chi^{(d)}$. We show that $\mathrm{Sw}_{P_i}(\chi) \leq n_i - 1$. To do this, it is enough to prove that the Swan conductor of $\chi^{(n_i)}$, with respect to the DVR at the generic point of the blow-up of $C^{(n_i)}$ along n_iP_i, is zero, by Corollary 2.9. We may assume that k is algebraically closed (hence $d_i = 1$). Note that the right vertical arrow p in (4-1) factors through $\tilde{C}_{\mathfrak{m}}^{(d)}$:

$$U^{(d)} \to \tilde{C}_{\mathfrak{m}}^{(d)} \to \mathrm{Pic}_{C,\mathfrak{m}}^{d}.$$

Since the pullback of ρ^d by p is $\chi^{(d)}$, we find that $\chi^{(d)}$ is unramified at the generic point of the complement $\tilde{C}_{\mathfrak{m}}^{(d)} \setminus U^{(d)}$. Thus, by Lemma 4.1 and Corollary 3.8, the character

$$\chi^{(n_1)} \boxtimes 1 \boxtimes \cdots \boxtimes 1 + 1 \boxtimes \chi^{(n_2)} \boxtimes \cdots \boxtimes 1 + \cdots + 1 \boxtimes \cdots \boxtimes \chi^{(d-\deg \mathfrak{m})}$$

on $U^{(n_1)} \times \cdots \times U^{(n_r)} \times U^{(d-\deg \mathfrak{m})}$ is unramified at the generic point of the exceptional divisor of the blow-up of $C^{(n_1)} \times \cdots \times C^{(n_r)} \times C^{(d-\deg \mathfrak{m})}$ along $\{n_1 P_1\} \times \cdots \times \{n_r P_r\} \times C^{(d-\deg \mathfrak{m})}$. Using Corollary 2.6 repeatedly, the assertion is proved.

Thus the map $B \to A$, pulling back by $U \to \mathrm{Pic}_{C,\mathfrak{m}}^1$, is well-defined. We denote this map by Ψ.

First we show the injectivity of Ψ. Take ρ from the kernel of Ψ. Since the multiplication map $\mathrm{Pic}_{C,\mathfrak{m}}^n \times \mathrm{Pic}_{C,\mathfrak{m}}^m \to \mathrm{Pic}_{C,\mathfrak{m}}^{n+m}$ and the two projections $\mathrm{Pic}_{C,\mathfrak{m}}^n \times \mathrm{Pic}_{C,\mathfrak{m}}^m \to \mathrm{Pic}_{C,\mathfrak{m}}^n, \mathrm{Pic}_{C,\mathfrak{m}}^m$ have geometrically connected fibers, the triviality of two of ρ^n, ρ^m, ρ^{n+m} implies the triviality of the other. Thus it is enough to show the triviality of ρ^d for sufficiently large d. Consider the diagram (4-1). By Lemma 4.2, we know that $p^*\rho^d$ is trivial, which implies that ρ^d is trivial by Corollary 3.9.

The surjectivity of Ψ is proved as follows. Take $\chi \in A$. Let d be an integer satisfying (3-1) for $\mathfrak{m}$. Corollary 2.9, Proposition 3.7, and Lemma 4.1 imply that the character $\chi^{(d)}$ extends to a character $\tilde{\chi}^{(d)}$ on $\tilde{C}_{\mathfrak{m}}^{(d)}$. Corollary 3.9 implies that $\tilde{\chi}^{(d)}$ descends to a character ρ^d on $\mathrm{Pic}_{C,\mathfrak{m}}^d$. Let d_1 and d_2 be integers which satisfy (3-1). The commutative diagram

$$\begin{array}{ccc} U^{(d_1)} \times U^{(d_2)} & \longrightarrow & U^{(d_1+d_2)} \\ \downarrow & & \downarrow \\ \mathrm{Pic}_{C,\mathfrak{m}}^{d_1} \times \mathrm{Pic}_{C,\mathfrak{m}}^{d_2} & \xrightarrow{q} & \mathrm{Pic}_{C,\mathfrak{m}}^{d_1+d_2} \end{array}$$

and the fact that the left vertical map has geometrically connected fibers show $q^*\rho^{d_1+d_2} = \rho^{d_1} \boxtimes 1 + 1 \boxtimes \rho^{d_2}$. Fix a nonzero effective Cartier divisor D on U such that $\deg D$ satisfies (3-1). Let ξ be the pullback of $\rho^{\deg D}$ by the map $\mathrm{Spec}(k) \to \mathrm{Pic}_{C,\mathfrak{m}}^{\deg D}$, corresponding to the point D. For an arbitrary integer n, take a natural number m so large that the integer $n + m \deg D$ satisfies (3-1). Define $\rho^n := f^*\rho^{n+m\deg D} \cdot a^*\xi^{-m}$, where $f : \mathrm{Pic}_{C,\mathfrak{m}}^n \to \mathrm{Pic}_{C,\mathfrak{m}}^{n+m\deg D}$ is multiplication by $\mathcal{O}_C(mD)$ and $a : \mathrm{Pic}_{C,\mathfrak{m}}^n \to \mathrm{Spec}(k)$ is the structure map. This construction does not depend on m, since the multiplicativity of ρ^n is already verified for large n. By the same reason, the characters ρ^n form a multiplicative element on $\mathrm{Pic}_{C,\mathfrak{m}}$. The equality $\chi = \Psi(\rho)$ follows from the commutative diagram

$$\begin{array}{ccccc} U & \xrightarrow{(\mathrm{id},g)} & U \times U^{(\deg D)} & \longrightarrow & U^{(\deg D+1)} \\ & & \downarrow & & \downarrow \\ & & \mathrm{Pic}_{C,\mathfrak{m}}^1 \times \mathrm{Pic}_{C,\mathfrak{m}}^{\deg D} & \longrightarrow & \mathrm{Pic}_{C,\mathfrak{m}}^{\deg D+1}, \end{array}$$

where g is the composition of the structure map $U \to \mathrm{Spec}(k)$ and the map $\mathrm{Spec}(k) \to U^{(\deg D)}$ corresponding to the divisor D. Indeed, the pullback of $\rho^{\deg D+1}$ by the map $U \to U \times U^{(\deg D)} \to U^{(\deg D+1)} \to \mathrm{Pic}^{\deg D+1}_{C,\mathfrak{m}}$ is $\chi \cdot b^*\xi$, where $b : U \to \mathrm{Spec}(k)$ is the structure map. On the other hand, the pullback of $\rho^{\deg D+1}$ the other way is $\Psi(\rho) \cdot b^*\xi$. □

To deduce Theorem 1.1 from Theorem 1.2, first we recall basic facts on torsors.

Assume that k is algebraically closed. Fix a connected commutative algebraic k-group G. Let $\mathcal{C}(G)$ be the category as follows. The objects are pairs $(H, \phi\colon H \to G)$ where H are connected commutative algebraic k-groups and ϕ are abelian isogenies. The morphisms $(H_1, \phi_1\colon H_1 \to G) \to (H_2, \phi_2\colon H_2 \to G)$ are pairs (f, g) where $f\colon H_1 \to H_2$ is a morphism of k-group schemes such that $\phi_2 \circ f = \phi_1$ and $g\colon H_1 \to H_2$ is a compatible morphism of torsors such that $\phi_2 \circ g = \phi_1$. Here we regard H_1 (resp. H_2) itself as an H_1-torsor (resp. H_2-torsor) by the multiplication. Note that the kernels of ϕ_i are constant k-schemes since H_i are Galois coverings of G.

Lemma 4.3. *Let the notation be as above. Let $(H_i, \phi_i\colon H_i \to G)$ be objects in $\mathcal{C}(G)$ for $i = 1, 2$.*

(1) *If there exists a morphism $H_1 \to H_2$ of G-*schemes, *there exists a unique morphism $f\colon H_1 \to H_2$ of k-group schemes with $\phi_2 \circ f = \phi_1$.*

(2) *The map*

$$\mathrm{Hom}((H_1, \phi_1\colon H_1 \to G), (H_2, \phi_2\colon H_2 \to G)) \to \mathrm{Hom}_G(H_1, H_2)$$

*sending $(f, g) \mapsto g$ is bijective. Here the target is the set of morphisms of G-*schemes.

Proof. Let $e_i \in H_i(k)$ be the units.

(1) Uniqueness follows from the fact that H_i are connected étale coverings of G and such an f must send e_1 to e_2. Let $f\colon H_1 \to H_2$ be the G-morphism which sends e_1 to e_2. Such an f does exist since H_2 is Galois over G. We need to show that the diagram

$$\begin{array}{ccc} H_1 \times H_1 & \xrightarrow{f \times f} & H_2 \times H_2 \\ \downarrow & & \downarrow \\ H_1 & \xrightarrow{f} & H_2, \end{array}$$

where the vertical maps are the multiplications, is commutative. This follows since $H_1 \times H_1$ is a connected étale covering of $G \times G$ and the two maps send (e_1, e_1) to e_2.

(2) Injectivity follows since a goup homomorphism $f\colon H_1 \to H_2$ over G is unique if it exists by (1). We show the surjectivity. Thus we assume that there is a group homomorphism $f\colon H_1 \to H_2$ over G. Let $g\colon H_1 \to H_2$ be a morphism of G-schemes. Since H_1 is a connected étale covering of G, this is uniquely determined by the image $a := g(e_1)$, which is contained in $\ker\phi_2$. Let $g'\colon H_1 \to H_2$ be the compatible morphism of torsors which sends e_1 to a. Since $a \in \ker\phi_2$ and $\phi_2 \circ f = \phi_1$, this is a morphism of G-schemes. Thus we have $g = g'$. □

Proof of Theorem 1.1. Let (G^0, G^1) be a connected abelian covering of $(\mathrm{Pic}^0_{C,\mathfrak{m}}, \mathrm{Pic}^1_{C,\mathfrak{m}})$. Since the d-th power of $\mathrm{Pic}^1_{C,\mathfrak{m}}$ is isomorphic to $\mathrm{Pic}^d_{C,\mathfrak{m}}$ as $\mathrm{Pic}^0_{C,\mathfrak{m}}$-torsors, the d-th power G^d of G^1 is naturally equipped with a compatible morphism $G^d \to \mathrm{Pic}^d_{C,\mathfrak{m}}$ of torsors. Let K be the kernel of the map $G^0 \to \mathrm{Pic}^0_{C,\mathfrak{m}}$. This is a finite constant group since $G^0 \to \mathrm{Pic}^0_{C,\mathfrak{m}}$ is a Galois isogeny. Take a nontrivial homomorphism $\chi : K \to \mathbb{Q}/\mathbb{Z}$. This defines characters $\rho^d \in \mathrm{H}^1(\mathrm{Pic}^d_{C,\mathfrak{m}}, \mathbb{Q}/\mathbb{Z})$ for all d. From the construction, they form a multiplicative element on $\mathrm{Pic}_{C,\mathfrak{m}}$. Theorem 1.2 implies that the pullback of ρ^1 by $U \to \mathrm{Pic}^1_{C,\mathfrak{m}}$ is nontrivial and its ramification is bounded by $\mathfrak{m}$, which shows the first part of Theorem 1.1.

Define the category $\mathcal{C}_1$ as the category of geometrically connected abelian coverings of U whose ramifications are bounded by $\mathfrak{m}$ and the category $\mathcal{C}_2$ as the category of connected abelian coverings of $(\mathrm{Pic}^0_{C,\mathfrak{m}}, \mathrm{Pic}^1_{C,\mathfrak{m}})$. We have constructed a functor $\Phi : \mathcal{C}_2 \to \mathcal{C}_1$. We show that this functor is an equivalence of categories. We only treat the case when k is algebraically closed. General case follows from this special case by using an argument of Galois descent.

Assume that k is algebraically closed. Let $\mathcal{C} := \mathcal{C}(\mathrm{Pic}^0_{C,\mathfrak{m}})$ be the category defined above. In this case, fixing a closed point P of U, $\mathcal{C}_2$ is isomorphic to $\mathcal{C}$ via the isomorphism $\mathrm{Pic}^0_{C,\mathfrak{m}} \to \mathrm{Pic}^1_{C,\mathfrak{m}}$ of torsors, tensoring $\mathcal{O}_C(P)$.

We show that the functor $\Phi' : \mathcal{C} \to \mathcal{C}_1$, pulling back by the morphism $U \to \mathrm{Pic}^0_{C,\mathfrak{m}}$ sending Q to $\mathcal{O}_C(Q-P)$ is an equivalence. Faithfulness is obvious since there only occur connected coverings. To show fullness, let $(G_i, G_i \to \mathrm{Pic}^0_{C,\mathfrak{m}})$ be elements of $\mathcal{C}$ for $i = 1, 2$ and let $V_i := \Phi'(G_i, G_i \to \mathrm{Pic}^0_{C,\mathfrak{m}})$. By Lemma 4.3(2) and faithfulness, it is enough to show that, if there is a map $V_1 \to V_2$, there is a map $G_1 \to G_2$. The kernel K_i of $G_i \to \mathrm{Pic}^0_{C,\mathfrak{m}}$ is canonically identified with the Galois group of $V_i \to U$. If there is a map $V_1 \to V_2$, there is a map of abelian groups $h : K_1 \to K_2$, which is independent of the choice of $V_1 \to V_2$. We show the commutativity of the diagram

$$\begin{array}{ccc} & \pi_1(\mathrm{Pic}^0_{C,\mathfrak{m}}) & \\ {}^{p_1}\swarrow & & \searrow^{p_2} \\ K_1 & \xrightarrow{\quad h \quad} & K_2 \end{array} \tag{4-2}$$

where the downward diagonals are the canonical surjections. Assume that there is an element $\sigma \in \pi_1(\mathrm{Pic}^0_{C,\mathfrak{m}})$ such that $p_2(\sigma) \neq hp_1(\sigma)$. Take a group homomorphism $\rho^0 : K_2 \to \mathbb{Q}/\mathbb{Z}$ such that the images of $p_2(\sigma)$ and $hp_1(\sigma)$ are different. Since the characters $\rho^0 p_2$ and $\rho^0 hp_1$ are multiplicative and are pulled back to the same character via the map $U \to \mathrm{Pic}^0_{C,\mathfrak{m}}$, they are the same character, a contradiction. Thus the diagram (4-2) is commutative, which implies that the quotient group $G_1/\ker h$ of G_1 is isomorphic to G_2.

For essential surjectivity, we argue as follows. Let $V \in \mathcal{C}_1$ be a connected cyclic covering of U. Take a character on U whose kernel corresponds to V. By Theorem 1.3, this character is the pullback of a multiplicative character ρ^0 on $\mathrm{Pic}^0_{C,\mathfrak{m}}$. Let G^0 be an étale covering of $\mathrm{Pic}^0_{C,\mathfrak{m}}$ corresponding to the kernel of ρ^0. We need to show that G^0 has a group structure. By the definition, G^0 is connected. From the multiplicativity of ρ^0, we know that there is a commutative diagram

$$\begin{array}{ccc} G^0 \times G^0 & \xrightarrow{m_G} & G^0 \\ \downarrow & & \downarrow \\ \mathrm{Pic}^0_{C,\mathfrak{m}} \times \mathrm{Pic}^0_{C,\mathfrak{m}} & \longrightarrow & \mathrm{Pic}^0_{C,\mathfrak{m}}. \end{array}$$

Let us denote the map m_G multiplicatively. Let F be the fiber of $G^0 \to \mathrm{Pic}^0_{C,\mathfrak{m}}$ over $1 \in \mathrm{Pic}^0_{C,\mathfrak{m}}$. For distinct points $y_1, y_2 \in F$, the multiplication from right by y_1 and y_2, $G^0 \to G^0$ are distinct. Indeed, Assume that $xy_1 = xy_2$ for all $x \in G^0$. Take a point x in F. The multiplication from left by x, $G^0 \to G^0$ is a $\mathrm{Pic}^0_{C,\mathfrak{m}}$-morphism and sends y_1 and y_2 to the same point, which implies that $y_1 = y_2$ since G^0 is a connected covering of $\mathrm{Pic}^0_{C,\mathfrak{m}}$.

Thus there exists an element $e \in F$ such that $xe = x$ for all $x \in G^0$. Next we show the commutativity of m_G. This follows from the fact that $G^0 \times G^0$ is a connected covering of $\mathrm{Pic}^0_{C,\mathfrak{m}} \times \mathrm{Pic}^0_{C,\mathfrak{m}}$ and that the maps $G^0 \times G^0 \to G^0$, $(x, y) \mapsto xy$ and $(x, y) \mapsto yx$ send (e, e) to the same point e. The associativity is proved in a similar way. Therefore it is verified that G^0 has a commutative group structure such that $G^0 \to \mathrm{Pic}^0_{C,\mathfrak{m}}$ is a group homomorphism, hence an abelian isogeny. It is easy to show that G^0 is pulled back to V. For a general V, use the fact that V is a connected component of the finite projective limit of cyclic connected coverings which are quotients of V. □

Acknowledgement

The author would like to express his sincere gratitude to his advisor, Professor Takeshi Saito, for a lot of advice and patient encouragement. The author thanks an anonymous referee for a careful reading and numerous helpful comments. This work was supported by the Program for Leading Graduate Schools, MEXT, Japan.

References

[Bosch et al. 1990] S. Bosch, W. Lütkebohmert, and M. Raynaud, *Néron models*, Ergebnisse der Mathematik (3) **21**, Springer, 1990. MR Zbl

[Brylinski 1983] J.-L. Brylinski, "Théorie du corps de classes de Kato et revêtements abéliens de surfaces", *Ann. Inst. Fourier (Grenoble)* **33**:3 (1983), 23–38. MR Zbl

[Guignard 2019] Q. Guignard, "On the ramified class field theory of relative curves", *Algebra Number Theory* **13**:6 (2019), 1299–1326.

[Kato 1989] K. Kato, "Swan conductors for characters of degree one in the imperfect residue field case", pp. 101–131 in *Algebraic K-theory and algebraic number theory* (Honolulu, 1987), edited by M. R. Stein and R. K. Dennis, Contemp. Math. **83**, Amer. Math. Soc., Providence, RI, 1989. MR Zbl

[Lang 1956] S. Lang, "Unramified class field theory over function fields in several variables", *Ann. of Math.* (2) **64** (1956), 285–325. MR Zbl

[Laumon 1990] G. Laumon, "Faisceaux automorphes liés aux séries d'Eisenstein", pp. 227–281 in *Automorphic forms, Shimura varieties, and L-functions, I* (Ann Arbor, MI, 1988), edited by L. Clozel and J. S. Milne, Perspect. Math. **10**, Academic Press, Boston, 1990. MR Zbl

[Leal 2018] I. Leal, "On ramification in transcendental extensions of local fields", *J. Algebra* **495** (2018), 15–50. MR Zbl

[Polishchuk 2003] A. Polishchuk, *Abelian varieties, theta functions and the Fourier transform*, Cambridge Tracts in Math. **153**, Cambridge Univ. Press, 2003. MR Zbl

[Rosenlicht 1954] M. Rosenlicht, "Generalized Jacobian varieties", *Ann. of Math.* (2) **59** (1954), 505–530. MR Zbl

[Serre 1988] J.-P. Serre, *Algebraic groups and class fields*, Graduate Texts in Math. **117**, Springer, 1988. MR Zbl

[SGA 1 1971] A. Grothendieck, *Revêtements étales et groupe fondamental* (Séminaire de Géométrie Algébrique du Bois Marie 1960–1961), Lecture Notes in Math. **224**, Springer, 1971. MR Zbl

[SGA 4_3 1973] M. Artin, A. Grothendieck, and J. L. Verdier, *Théorie des topos et cohomologie étale des schémas, Tome 3: Exposés IX–XIX* (Séminaire de Géométrie Algébrique du Bois Marie 1963–1964), Lecture Notes in Math. **305**, Springer, 1973. MR Zbl

Communicated by Bjorn Poonen
Received 2018-05-20 Revised 2019-02-19 Accepted 2019-03-25

daichi@ms.u-tokyo.ac.jp *Graduate School of Mathematical Sciences, The University of Tokyo, Japan*

ALGEBRA AND NUMBER THEORY 13:6 (2019)
dx.doi.org/10.2140/ant.2019.13.1353

Algebraic monodromy groups of l-adic representations of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$

Shiang Tang

In this paper we prove that for any connected reductive algebraic group G and a large enough prime l, there are continuous homomorphisms

$$\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to G(\bar{\mathbb{Q}}_l)$$

with Zariski-dense image, in particular we produce the first such examples for SL_n, Sp_{2n}, Spin_n, E_6^{sc} and E_7^{sc}. To do this, we start with a mod-l representation of $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ related to the Weyl group of G and use a variation of Stefan Patrikis' generalization of a method of Ravi Ramakrishna to deform it to characteristic zero.

1. Introduction

For a split connected reductive group G and a prime number l, it is natural to study two types of continuous representations of $\Gamma_{\mathbb{Q}} = \mathrm{Gal}\,(\bar{\mathbb{Q}}/\mathbb{Q})$: the mod l representations

$$\bar{\rho} : \Gamma_{\mathbb{Q}} \to G(\bar{\mathbb{F}}_l)$$

and the l-adic representations

$$\rho : \Gamma_{\mathbb{Q}} \to G(\bar{\mathbb{Q}}_l)$$

where we use the discrete topology for $G(\bar{\mathbb{F}}_l)$ and the l-adic topology for $G(\bar{\mathbb{Q}}_l)$. Mod l representations of $\Gamma_{\mathbb{Q}}$ are closely related to the inverse Galois problem for finite groups of Lie type, which asks for the existence of surjective homomorphisms $\bar{\rho} : \Gamma_{\mathbb{Q}} \twoheadrightarrow G(k)$ for k a finite extension of $\mathbb{F}_l$. It is still wide open, even for small groups such as SL_2. If we replace $\Gamma_{\mathbb{Q}}$ by Γ_F for some number field F, it is not hard to show that every finite group is a Galois group over *some* number field, but if we insist on $\Gamma_{\mathbb{Q}}$ then the problem becomes very difficult. On the other hand, we can ask for its analogs in the l-adic world:

Question 1. Are there continuous homomorphisms $\rho : \Gamma_{\mathbb{Q}} \to G(\bar{\mathbb{Q}}_l)$ with Zariski-dense image?

We also ask a refined question which takes geometric Galois representations (in the sense of [Fontaine and Mazur 1995]) into account:

Question 2. Are there continuous geometric Galois representations $\rho : \Gamma_{\mathbb{Q}} \to G(\bar{\mathbb{Q}}_l)$ with Zariski-dense image?

MSC2010: 11F80.
Keywords: Galois representation, Galois deformation theory.

This paper gives a complete answer to Question 1. We shall call a reductive group G an l-adic algebraic monodromy group, or simply an l-adic monodromy group for $\Gamma_{\mathbb{Q}}$ if the homomorphisms in Question 1 exist, and a geometric l-adic monodromy group for $\Gamma_{\mathbb{Q}}$ if the homomorphisms in Question 2 exist. We prove the following theorem which gives an almost complete answer to Question 1:

Theorem 1.1 (Main Theorem). *Let G be a connected reductive algebraic group. Then there are continuous homomorphisms*

$$\rho_l : \Gamma_{\mathbb{Q}} \to G(\bar{\mathbb{Q}}_l)$$

with Zariski-dense image for large enough primes l.

The key cases of our main theorem are contained the following theorem:

Theorem 1.2. *For a simple algebraic group G, there are infinitely many continuous homomorphisms*

$$\rho_l : \Gamma_{\mathbb{Q}} \to G(\bar{\mathbb{Q}}_l)$$

with Zariski-dense image for l large enough. We impose the condition $l \equiv 1(4)$ for $G = B_n^{\mathrm{sc}}, C_n^{\mathrm{sc}}$, and impose $l \equiv 1(3)$ for $G = E_7^{\mathrm{sc}}$.

Remark 1.3. Patrikis has shown that $E_7^{\mathrm{ad}}, E_8, F_4, G_2$ and the L-group of an outer form of E_6^{ad} are geometric l-adic monodromy groups for $\Gamma_{\mathbb{Q}}$, so we will not discuss these cases in the proof of the above theorem. The congruence conditions on l for $G = B_n^{\mathrm{sc}}, C_n^{\mathrm{sc}}, E_7^{\mathrm{sc}}$ can be removed using a theorem proved by Fakhruddin, Khare and Patrikis [2018]. We record it in Theorem 3.21.

It is shown in [Cornut and Ray 2018] that for sufficiently large regular primes p (i.e., a prime p that does not divide the class number of $\mathbb{Q}(\mu_p)$) and for a simple, adjoint group G, there exist a continuous representation of $\Gamma_{\mathbb{Q}}$ into $G(\mathbb{Q}_p)$ with image between the pro-p and the standard Iwahori subgroups of G, which generalizes a theorem of Greenberg [2016] for GL_n. In particular, the image of the Galois group is Zariski-dense. Their construction is nongeometric and is very different from ours. It is unknown whether or not there are infinitely many regular primes, however.

It is an interesting question whether (for instance) SL_n can be a geometric monodromy group for $\Gamma_{\mathbb{Q}}$. The following example shows that Question 2 is more subtle than Question 1, and we should not expect an answer as clean as Theorem 1.1.

Example 1.4. Assuming the Fontaine–Mazur and the Langlands conjectures (see [Fontaine and Mazur 1995; Buzzard and Gee 2014]), there is no homomorphism $\rho : \Gamma_{\mathbb{Q}} \to \mathrm{SL}_2(\bar{\mathbb{Q}}_l)$ that is unramified almost everywhere, potentially semistable at l, and has Zariski-dense image.

Proof. In fact, by the Fontaine–Mazur and the Langlands conjectures, if such ρ exists, then $\rho = \rho_\pi$ for some cuspidal automorphic representation π on $\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q}})$. But ρ is even, i.e., $\det \rho(c) = 1$, π_∞ (the archimedean component of π) is a principal series representation, and π is associated to a Maass form. Therefore, by the Fontaine–Mazur conjecture, ρ_π has finite image, a contradiction. □

In contrast, Theorem 1.2 shows in particular that SL_2 is an l-adic monodromy group for $\Gamma_{\mathbb{Q}}$. On the other hand, SL_2 can be a geometric l-adic monodromy group for Γ_F for *some* finite extension $F/\mathbb{Q}$.

Example 1.5. Let f be a non-CM new eigenform of weight 3, level N, with a nontrivial nebentypus character ε. Such f exist for suitable N, see [LMFDB 2013]. We write E for the field of coefficients of f. Then for all l and $\lambda \mid l$, there is a continuous representation $r_{f,\lambda} : \Gamma_{\mathbb{Q}} \to \mathrm{GL}_2(E_\lambda)$ which is unramified outside $\{v : v \mid Nl\}$ and $\mathrm{Tr}(r_{f,\lambda}(\mathrm{Fr}_p)) = a_p$ for p not dividing Nl, with a_p the p-th Hecke eigenvalue of f. We have $\det(r_{f,\lambda}) = \kappa^2\varepsilon$ where κ is the l-adic cyclotomic character. By a theorem of Ribet [1985, Theorem 2.1], for almost all l, $\bar{r}_{f,\lambda}(\Gamma_{\mathbb{Q}})$ contains $\mathrm{SL}_2(k)$ for a subfield k of k_λ (the residue field of E_λ). It follows that $r_{f,\lambda}$ has Zariski dense image. If we let F be a finite extension of $\mathbb{Q}$ that trivializes ε, then the image of $r' := \kappa^{-1} \cdot r_{f,\lambda}|_{\Gamma_F}$ lands in $\mathrm{SL}_2(E_\lambda)$ and is Zariski-dense.

The classical groups GSp_n, GSO_n are known as geometric l-adic monodromy groups. Recent work of Arno Kret and Sug Woo Shin [2016] obtains GSpin_{2n+1} as a geometric l-adic monodromy group and Nick Katz [2018] constructs geometric Galois representations with monodromy group GL_n. On the other hand, most of the exceptional algebraic groups are known as geometric l-adic monodromy groups, established in the work of Dettweiler and Reiter [2010], Zhiwei Yun [2014] and Stefan Patrikis [2016]. Patrikis [2016] constructs geometric Galois representations for $\Gamma_{\mathbb{Q}}$ with full algebraic monodromy groups for essentially all exceptional groups of adjoint type. Along the way, Patrikis has obtained an extension to general reductive groups of Ravi Ramakrishna's techniques for lifting odd two-dimensional Galois representations to geometric l-adic representations in [Ramakrishna 2002].

For the rest of this section, we sketch the strategy for proving Theorem 1.2, which makes use of Patrikis' generalization of Ramakrishna's techniques but is very different from his arguments in many ways. For the rest of this section, we assume that G is a simple algebraic group defined over $\mathbb{Z}_l$ with a split maximal torus T. Let $\Phi = \Phi(G, T)$ be the associated root system. Let $\mathcal{O}$ be the ring of integers of an extension of $\mathbb{Q}_l$ whose reduction modulo its maximal ideal is isomorphic to k, a finite extension of $\mathbb{F}_l$. We start with a well-chosen $\bmod l$ representation and then use a variant of Ramakrishna's method to deform it to characteristic zero with big image. Achieving this is a balancing act between two difficulties: the inverse Galois problem for $G(k)$ is difficult, so we want the residual image to be relatively "small"; on the other hand, Ramakrishna's method works when the residual image is "big".

Let us recall a construction used in [Patrikis 2016]. Patrikis uses the principal GL_2 homomorphism to construct the residual representation

$$\bar{\rho} : \Gamma_{\mathbb{Q}} \xrightarrow{\bar{r}} \mathrm{GL}_2(k) \xrightarrow{\varphi} G(k)$$

for $\bar{r}$ a surjective homomorphism constructed from modular forms and φ a principal GL_2 homomorphism (for its definition, see [Serre 1996; Patrikis 2016, Section 7.1]). But the principal GL_2 is defined only when $\rho^\vee$ (the half-sum of coroots) is in the cocharacter lattice $X_*(T)$, which is not the case for $G = \mathrm{SL}_{2n}$, Sp_{2n}, E_7^{sc}, etc. On the other hand, the principal SL_2 is always defined but it is not known whether $\mathrm{SL}_2(k)$ is a Galois group over $\mathbb{Q}$ and the surjectivity of $\bar{r}$ is crucial in applying Ramakrishna's method.

For this reason, we use a different construction. To simplify notation, we use G, T to denote $G(k)$, $T(k)$, respectively. We consider the following exact sequence of finite groups, which we shall refer to as the N-T sequence:

$$1 \to T \to N_G(T) \xrightarrow{\pi} \mathcal{W} \to 1$$

where $\mathcal{W} = N_G(T)/T$ is the Weyl group of G. We want to take $N = N_G(T)$ as the image of the residual representation $\bar{\rho}$. It turns out that the adjoint action of N on the Lie algebra $\mathfrak{g}$ over k decomposes into at most three irreducible pieces (Corollary 2.2), which is very good for applying Ramakrishna's techniques. It has been known for a long time that $\mathcal{W}$ is a Galois group over $\mathbb{Q}$, but what we need is to realize N as a Galois group over $\mathbb{Q}$. A natural approach would be solving the embedding problem posed by the N-T sequence, i.e., to suppose there is a Galois extension $K/\mathbb{Q}$ realizing $\mathcal{W}$, and then to find a finite Galois extension $K'/\mathbb{Q}$ containing K such that the natural surjective homomorphism $\mathrm{Gal}(K'/\mathbb{Q}) \twoheadrightarrow \mathrm{Gal}(K/\mathbb{Q})$ realizes $\pi : N \twoheadrightarrow \mathcal{W}$.

This embedding problem is solvable when the sequence *splits*, by an elementary case of a famous theorem of Igor Shafarevich, see [Serre 1992, Claim 2.2.5]. In [Adams and He 2017], the splitting of the N-T sequence is determined completely; for instance, it does not split for $G = \mathrm{SL}_n$, Sp_{2n}, Spin_n, E_7. We find our way out by replacing N with a suitable subgroup N' for which the decomposition of the adjoint representation remains the same, then realizing N' as a Galois group over $\mathbb{Q}$ with certain properties, see Sections 2A3–2A6. Finally, we define our residual representation $\bar{\rho}$ to be the composite

$$\Gamma_{\mathbb{Q}} \twoheadrightarrow N' \to G = G(k)$$

where the first arrow comes from the realization of N' as a Galois group over $\mathbb{Q}$ and the second arrow is the inclusion map. We write $\bar{\rho}(\mathfrak{g})$ for the Lie algebra $\mathfrak{g}/k$ equipped with a $\Gamma_{\mathbb{Q}}$-action induced by the homomorphism

$$\Gamma_{\mathbb{Q}} \xrightarrow{\bar{\rho}} G \xrightarrow{\mathrm{Ad}} \mathrm{GL}(\mathfrak{g}).$$

Now we explain how to deform $\bar{\rho}$ to characteristic zero. This is the hardest part. For a residual representation

$$\bar{\rho} : \Gamma_{\mathbb{Q}} \to G(k)$$

unramified outside a finite set of places S containing the archimedean place and a global deformation condition for $\bar{\rho}$ (which consists of a local deformation condition for each $v \in S$), a typical question in Galois deformation theory is to find continuous l-adic lifts

$$\rho : \Gamma_{\mathbb{Q}} \to G(\mathcal{O})$$

of $\bar{\rho}$ such that for all v, $\rho|_{\Gamma_{\mathbb{Q}_v}}$ (we fix an embedding $\bar{\mathbb{Q}} \to \bar{\mathbb{Q}}_v$) satisfies the prescribed local deformation condition at v. If this can be done, then we can make the image of ρ Zariski-dense in $G(\bar{\mathbb{Q}}_l)$ by specifying a certain type of local deformation condition at a suitable unramified prime.

Ravi Ramakrishna [2002] has an ingenious method for obtaining the desired lifts, which has been generalized and axiomatized in [Taylor 2003; Clozel et al. 2008; Patrikis 2016] and others. By the Poitou–Tate exact sequence, if the dual Selmer group

$$H^1_{\mathcal{L}^\perp}(\Gamma_{\mathbb{Q},S}, \bar{\rho}(\mathfrak{g})(1))$$

associated to the global deformation condition for $\bar{\rho}$ vanishes, then such lifts exist. Here $\mathcal{L}$ and $\mathcal{L}^\perp$ are the Selmer system and dual Selmer system of tangent spaces and annihilators of the tangent spaces under the Tate pairing of the given global deformation condition, respectively. Ramakrishna discovered that if one imposes additional local deformation conditions of "Ramakrishna type" in place of the unramified ones at a finite set of well-chosen places of $\mathbb{Q}$ disjoint from S, then the new dual Selmer group will vanish. However, this technique is very sensitive to the image of $\bar{\rho}$, which has to be "big" to make things work; if $\bar{\rho}(\mathfrak{g})$ is irreducible then all is good, but finding such a $\bar{\rho}$ can be very difficult. In practice, we would prefer those $\bar{\rho}$ for which $\bar{\rho}(\mathfrak{g})$ does not decompose too much. Unfortunately, the form of Ramakrishna's method in [Patrikis 2016] (see Theorem 3.4 and its proof for an account of this) does not work for our $\bar{\rho}$.

Inspired by the use of Ramakrishna's method in [Clozel et al. 2008], we surmount this by making two observations. For our $\bar{\rho}$, $\bar{\rho}(\mathfrak{g})$ decomposes into $\bar{\rho}(\mathfrak{t})$ (the Lie algebra of T over k equipped with an irreducible action of $\bar{\rho}(\Gamma_{\mathbb{Q}})$) and a complement (see Corollary 2.2). Our *first observation* is that if

$$H^1_{\mathcal{L}}(\Gamma_{\mathbb{Q},S}, \bar{\rho}(\mathfrak{t}))$$

(see Definition 3.7) vanishes, then we can kill the full dual Selmer group using Ramakrishna's method; moreover, we *cannot* find an auxiliary prime $w \notin S$ at which the Ramakrishna deformation condition (see Definition 3.1) satisfies

$$h^1_{\mathcal{L}^\perp \cup L_w^{\mathrm{Ram},\perp}}(\Gamma_{\mathbb{Q},S\cup w}, \bar{\rho}(\mathfrak{g})(1)) < h^1_{\mathcal{L}^\perp}(\Gamma_{\mathbb{Q},S}, \bar{\rho}(\mathfrak{g})(1))$$

when $H^1_{\mathcal{L}}(\Gamma_{\mathbb{Q},S}, \bar{\rho}(\mathfrak{t})) \neq \{0\}$. But it is hard to achieve $H^1_{\mathcal{L}}(\Gamma_{\mathbb{Q},S}, \bar{\rho}(\mathfrak{t})) = \{0\}$ in general.

Our *second observation* is as follows: suppose that $0 \neq h^1_{\mathcal{L}}(\Gamma_{\mathbb{Q},S}, \bar{\rho}(\mathfrak{t})) \leq h^1_{\mathcal{L}^\perp}(\Gamma_{\mathbb{Q},S}, \bar{\rho}(\mathfrak{t})(1))$ (the inequality is easy to guarantee), and let ϕ be a nontrivial class in $H^1_{\mathcal{L}^\perp}(\Gamma_{\mathbb{Q},S}, \bar{\rho}(\mathfrak{t})(1))$. We can then find an auxiliary prime $w \notin S$ with a Ramakrishna deformation L_w^{Ram} such that $\phi|_{\Gamma_{\mathbb{Q}_w}} \notin L_w^{\mathrm{Ram},\perp}$, which implies

$$h^1_{\mathcal{L}^\perp \cup L_w^{\mathrm{Ram},\perp}}(\Gamma_{\mathbb{Q},S\cup w}, \bar{\rho}(\mathfrak{t})(1)) < h^1_{\mathcal{L}^\perp \cup L_w^\perp}(\Gamma_{\mathbb{Q},S\cup w}, \bar{\rho}(\mathfrak{t})(1)),$$

where L_w is the intersection of L_w^{Ram} and the unramified condition at w. It turns out that (see the proof of Proposition 3.13) the right side of the inequality equals $h^1_{\mathcal{L}^\perp}(\Gamma_{\mathbb{Q},S}, \bar{\rho}(\mathfrak{t})(1))$; then a double invocation of Wiles' formula gives

$$h^1_{\mathcal{L} \cup L_w^{\mathrm{Ram}}}(\Gamma_{\mathbb{Q},S\cup w}, \bar{\rho}(\mathfrak{t})) < h^1_{\mathcal{L}}(\Gamma_{\mathbb{Q},S}, \bar{\rho}(\mathfrak{t})).$$

By induction, we can enlarge $\mathcal{L}$ finitely many times to make $H^1_{\mathcal{L}}(\Gamma_{\mathbb{Q},S}, \bar{\rho}(\mathfrak{t}))$ vanish, which in turn allows us (see the proof of Theorem 3.16) to enlarge $\mathcal{L}$ even further to make $H^1_{\mathcal{L}^\perp}(\Gamma_{\mathbb{Q},S}, \bar{\rho}(\mathfrak{g})(1))$ vanish, as remarked in the first observation. Thus we obtain an l-adic lift $\rho : \Gamma_{\mathbb{Q}} \to G(\mathcal{O})$ satisfying the prescribed local deformation conditions.

The above variation of Ramakrishna's method is devised very specifically for the residual representation we construct. We do not know how to generalize these ideas to the case of an arbitrary residual representation.

Remark 1.6. For technical reasons, our method does not work for SL_2, SL_3, Spin_7; see Proposition 2.9 and Remark 2.14. Nevertheless, an easy variant of Ramakrishna's original method applies to SL_2, Patrikis' extension of Ramakrishna's method applies to SL_3 and Spin_7. Patrikis' method also applies to E_6^{sc} with minor modifications, and we use it in this paper. Our method should work for E_6^{sc} as well, modulo an instance of the inverse Galois theory, but we do not pursue it here.

Notation. For a field F (typically $\mathbb{Q}$ or $\mathbb{Q}_p$), we let Γ_F denote $\mathrm{Gal}(\bar{F}/F)$ for some fixed choice of algebraic closure of $\bar{F}$ of F. When F is a number field, for each place v of F we fix once and for all embeddings $\bar{F} \to \bar{F}_v$, giving rise to inclusions $\Gamma_{F_v} \to \Gamma_F$. If S is a finite set of places of F. we let $\Gamma_{F,S}$ denote $\mathrm{Gal}(F_S/F)$, where F_S is the maximal extension of F in $\bar{F}$ unramified outside of S. If v is a place of F outside S, we write Fr_v for the corresponding arithmetic frobenius element in $\Gamma_{F,S}$. When $F = \mathbb{Q}$, we will sometimes write Γ_v for Γ_{F_v} and Γ_S for $\Gamma_{\mathbb{Q},S}$. For a representation ρ of Γ_F, we let $F(\rho)$ denote the fixed field of $\mathrm{Ker}(\rho)$.

Consider a group Γ, a ring A, an algebraic group G over $\mathrm{Spec}(A)$, and a homomorphism $\rho : \Gamma \to G(A)$. We write $\mathfrak{g}$ for both the Lie algebra of G and the $A[G]$-module induced by the adjoint action. We let $\rho(\mathfrak{g})$ denote the $A[\Gamma]$-module with underlying A-module $\mathfrak{g}$ induced by ρ. Similarly, for a $A[G]$-submodule M of $\mathfrak{g}$, we write $\rho(M)$ for the $A[\Gamma]$-module with underlying A-module M induced by ρ.

We call an algebraic group *simple* if it is connected, nonabelian and has no proper normal algebraic subgroups except for finite subgroups. It is sometimes called an almost simple group in the literature. Consider a simple algebraic group G, we write G^{sc} and G^{ad} for the simply connected form and adjoint form of G, respectively.

Let $\mathcal{O}$ be the ring of integers of a finite extension of $\mathbb{Q}_l$. We let $\mathrm{CNL}_{\mathcal{O}}$ denote the category of complete noetherian local $\mathcal{O}$-algebras for which the structure map $\mathcal{O} \to R$ induces an isomorphism on residue fields.

All the Galois cohomology groups we consider will be k-vector spaces for k a finite extension of $\mathbb{F}_l$. We abbreviate $\dim H^n(-)$ by $h^n(-)$.

We write κ for the l-adic cyclotomic character, and $\bar{\kappa}$ for its $\bmod l$ reduction.

2. Constructions of residual representations

In this section, we construct residual representations

$$\bar{\rho} : \Gamma_{\mathbb{Q}} \to G(\bar{\mathbb{F}}_l)$$

for G a simple, *simply connected* algebraic group.

2A. ***Constructions based on the Weyl groups.*** Let k be a finite extension of $\mathbb{F}_l$. We consider the group of k-points of the normalizer of a split maximal torus of G and hope to realize it as the Galois group of some extension of $\mathbb{Q}$.

2A1. *Some group-theoretic results.* We recall a property of the Weyl group of an *irreducible* root system Φ.

Lemma 2.1. *The Weyl group $\mathcal{W}$ acts irreducibly on the $\mathbb{C}$-vector space spanned by Φ and transitively on roots of the same length.*

Let $G = G(k)$ and $T = T(k)$, a maximal split torus of G. Let $\Phi = \Phi(G, T)$ and $N = N_G(T)$.

Corollary 2.2. *For any $\alpha, \beta \in \Phi$ of the same length, there exists $w \in N$ such that $\mathrm{Ad}(w)\mathfrak{g}_\alpha = \mathfrak{g}_\beta$. The adjoint action $\mathrm{Ad}(N)$ on $\mathfrak{g}$ decomposes into submodules $\mathfrak{t}$ and*

$$\mathfrak{g}_\Phi := \sum_{\alpha\in\Phi} \mathfrak{g}_\alpha$$

when Φ is simply laced, and is the direct sum of $\mathfrak{t}$,

$$\mathfrak{g}_l := \sum_{\alpha\in\Phi,\ \alpha \text{ is long}} \mathfrak{g}_\alpha,$$

and

$$\mathfrak{g}_s := \sum_{\alpha\in\Phi,\ \alpha \text{ is short}} \mathfrak{g}_\alpha$$

otherwise.

Moreover, as an N-module, $\mathfrak{t}$ is irreducible, and $\mathfrak{g}_\Phi, \mathfrak{g}_l, \mathfrak{g}_s$ are irreducible if l is sufficiently large.

Proof. It suffices to show that $\mathfrak{g}_\Phi, \mathfrak{g}_l, \mathfrak{g}_s$ are irreducible N-modules. We will only show that $\mathfrak{g}_\Phi$ is irreducible, for the other two cases are similar. Take a nonzero vector $X \in \mathfrak{g}_\Phi$, write $X = \sum_{1\le i\le k} X_i$ where $0 \ne X_i \in \mathfrak{g}_{\alpha_i}$ for some distinct roots $\alpha_1, \alpha_2, \dots, \alpha_k \in \Phi$. Since l is sufficiently large, we can choose $t \in T$ such that $\alpha(t)$, $\alpha \in \Phi$ are all distinct. We have

$$\mathrm{ad}(t^j)X = \sum_i \alpha_i(t)^j X_i$$

where $0 \le j \le k-1$. As the X_i's are linearly independent and the determinant of the coefficient matrix is nonzero,

$$X, \mathrm{ad}(t)X, \mathrm{ad}(t^2)X, \dots, \mathrm{ad}(t^{k-1})X$$

are linearly independent and hence they span the same subspace as $X_1, \dots, X_k$ do. In particular, $X_i \in \mathfrak{g}_{\alpha_i}$ belongs to the N-submodule of $\mathfrak{g}_\Phi$ generated by X. Because N acts transitively on the set of root spaces, it follows that X generates $\mathfrak{g}_\Phi$. Therefore, $\mathfrak{g}_\Phi$ is irreducible. □

Remark 2.3. Corollary 2.2 remains valid for a subgroup N' of N that maps onto a subgroup W' of W acting transitively on roots of the same length.

2A2. *Some results in the inverse Galois theory.* We record some elementary results (with proofs) about inverse Galois theory, some of which are modified in order to satisfy our purposes. See Serre's lecture notes [1992] for details.

Theorem 2.4. *For $n \geq 2$, there are infinitely many polynomials with $\mathbb{Q}$-coefficients that realize the symmetric group of n letters S_n (or the alternating group of n letters A_n) as a Galois group over $\mathbb{Q}$. Moreover, for S_n (or for A_n with $n \geq 4$), the polynomial can be chosen to have at least a pair of nonreal roots.*

Proof. See [Serre 1992, Sections 4.4 and 4.5]. For the last part, we consider the polynomial $f(X, T)$ on page 42 of [loc. cit.]. For any rational value of T, it has at most three real roots by inspection, so it must have at least a nonreal root when $n \geq 4$. For $n = 2, 3$, it is easy to find such polynomials. The demonstration for A_n is similar, see the polynomial $h(X, T)$ on page 44. □

The next result is an elementary case of a theorem of Igor Shafarevich, which will be used frequently in our constructions:

Theorem 2.5. *Let G be a finite group. Suppose that there is a finite Galois extension $K/\mathbb{Q}$ such that* $\mathrm{Gal}(K/\mathbb{Q}) \cong G$. *Let H be a finite abelian group with exponent m. Suppose that there is a* split *exact sequence of finite groups*

$$1 \to H \to S \to G \to 1.$$

Then there is a finite Galois extension $M/\mathbb{Q}$ containing K such that the natural surjective homomorphism $\mathrm{Gal}(M/\mathbb{Q}) \twoheadrightarrow \mathrm{Gal}(K/\mathbb{Q})$ *realizes the surjective homomorphism $S \twoheadrightarrow G$. In other words, a split embedding problem with abelian kernel always has a proper solution.*

Moreover, for any prime l that is outside the ramification locus of $K/\mathbb{Q}$ and prime to m, we can choose M so that l is unramified in M.

Proof. The argument is a minor modification of the proof of [Serre 1992, Claim 2.2.5]. Put $L = K(\mu_m)$. H can be regarded as a m-torsion module on which $\mathrm{Gal}\,(L/\mathbb{Q})$ acts. So there is a finite free $(\mathbb{Z}/m\mathbb{Z})[\mathrm{Gal}\,(L/\mathbb{Q})]$-module F of which H is a quotient. Suppose that r is the number of copies of $(\mathbb{Z}/m\mathbb{Z})[\mathrm{Gal}\,(L/\mathbb{Q})]$ in F. Let S' be the semidirect product of $\mathrm{Gal}\,(L/\mathbb{Q})$ and F. To solve the embedding problem posed by $1 \to H \to S \to G \to 1$, it suffices to solve the embedding problem posed by

$$1 \to F \to S' \to \mathrm{Gal}(L/\mathbb{Q}) \to 1.$$

Claim. There is a Galois extension $M'/\mathbb{Q}$ that solves the above embedding problem. Moreover, for any prime l that does not divide m and is outside the ramification locus of $K/\mathbb{Q}$, we can choose M' so that l is unramified in M'.

To see this, we choose places $v_1, \dots, v_r$ of $\mathbb{Q}$ away from l such that v_i *splits completely* in L. Let w_i be a place of L extending v_i, $1 \leq i \leq r$. Any place of L extending v_i can be written *uniquely* as σw_i for some $\sigma \in \mathrm{Gal}\,(L/\mathbb{Q})$. Let w_0 be a place of L extending l. For $1 \leq i \leq r$, choose $\theta_i \in \mathcal{O}_L$ such that for

$0 \le j \le r$, $(\sigma w_j)(\theta_i) = 1$ if $\sigma = 1$ and $i = j$ and 0 if not. The existence of θ_i follows from the weak approximation theorem. Let

$$M' = L(\sqrt[m]{\sigma\theta_i} \mid \sigma \in \mathrm{Gal}\,(L/\mathbb{Q}),\ 1 \le i \le r),$$

which is Galois over $\mathbb{Q}$, being the composite of L and the splitting field of the polynomial

$$\prod_i \prod_\sigma (T^m - \sigma\theta_i) \in \mathbb{Q}[T].$$

It is easy to see that $\mathrm{Gal}\,(M'/L)$ is isomorphic to F as $\mathrm{Gal}\,(L/\mathbb{Q})$-modules. In fact, for each i there is an isomorphism

$$\phi_i : \mathrm{Gal}(L(\sqrt[m]{\sigma\theta_i} \mid \sigma \in \mathrm{Gal}\,(L/\mathbb{Q}))/L) \cong (\mathbb{Z}/m\mathbb{Z})[\mathrm{Gal}(L/\mathbb{Q})].$$

For an element g on the left side and any $\sigma \in \mathrm{Gal}\,(L/\mathbb{Q})$, $g(\sqrt[m]{\sigma\theta_i}) = \zeta_\sigma \cdot \sqrt[m]{\sigma\theta_i}$ for some $\zeta_\sigma \in \mu_m \cong \mathbb{Z}/m\mathbb{Z}$. We then define

$$\phi_i(g) = \sum_\sigma \zeta_\sigma \cdot \sigma \in (\mathbb{Z}/m\mathbb{Z})[\mathrm{Gal}\,(L/\mathbb{Q})].$$

It is clear that ϕ_i is an isomorphism by our choice of θ_i. It follows that $\mathrm{Gal}\,(M'/L) \cong F$ by linear disjointness.

Therefore, we obtain an exact sequence

$$1 \to F \to \mathrm{Gal}(M'/\mathbb{Q}) \to \mathrm{Gal}(L/\mathbb{Q}) \to 1.$$

Since $F \cong \mathrm{Ind}_{\{1\}}^{\mathrm{Gal}\,(L/\mathbb{Q})} \mathbb{Z}/m\mathbb{Z}$, by Shapiro's lemma, $H^2(\mathrm{Gal}\,(L/\mathbb{Q}), F) = H^2(\{1\}, \mathbb{Z}/m\mathbb{Z}) = \{0\}$, hence the sequence splits. Thus, $\mathrm{Gal}(M'/\mathbb{Q}) \cong S'$.

It remains to show that l is unramified in M'. For any σ in $\mathrm{Gal}\,(L/\mathbb{Q})$ and for any i, w_0 is unramified in $L(\sqrt[m]{\sigma\theta_i})$, because $w_0(\sigma\theta_i) = 0$ and l does not divide m. So w_0 is unramified in their composite M'. On the other hand, l is unramified in K by assumption and is unramified in $\mathbb{Q}(\mu_m)$ since l does not divide m, so l is unramified in L. It follows that l is unramified in M', proves the claim.

Finally, letting M be the fixed field of the kernel of the natural surjective homomorphism $S' \twoheadrightarrow S$, we obtain a solution to the original embedding problem. □

2A3. SL_n. Let $G = \mathrm{SL}_n(k)$, so $\mathcal{W} \cong S_n$. By [Adams and He 2017], the N-T sequence splits only when n is odd. We consider the subgroup $\mathcal{W}' = A_n$ of $\mathcal{W}$. Let T be the maximal torus of diagonal elements in $\mathrm{SL}_n(k)$ and let Φ be $\Phi(G, T)$.

Lemma 2.6. *Suppose $n \ge 4$. Then A_n, as a subgroup of $\mathcal{W}$, acts transitively on Φ.*

Proof. This follows from the fact that A_n acts doubly transitively on $\{1, 2, \ldots, n\}$ if and only if $n \ge 4$. □

Let $N' = \pi^{-1}(\mathcal{W}')$, where π is the natural map from $N_G(T)$ to $\mathcal{W}$.

Lemma 2.7. *The following exact sequence of finite groups splits*:

$$1 \to T \to N' \to \mathcal{W}' \to 1.$$

Proof. We think of A_n as a subgroup of $N = N_G(T)$ by realizing it as the group of $n \times n$ *even* permutation matrices. Then A_n normalizes T and N' is a semidirect product of them. $\square$

Let l be large enough. Since T is abelian of exponent $|k| - 1$ which is prime to l, by Theorems 2.4 and 2.5, there is a surjection $\Gamma_{\mathbb{Q}} \twoheadrightarrow N'$ which is unramified at l. Moreover, by choosing a rational polynomial with nonreal roots that realizes A_n as a Galois group over $\mathbb{Q}$, we can make the complex conjugation map to an element away from the center of G. Define $\bar{\rho}$ to be the composite

$$\Gamma_{\mathbb{Q}} \twoheadrightarrow N' \xrightarrow{i} \mathrm{SL}_n(k).$$

Remark 2.8. It is trickier to realize N as a Galois group over $\mathbb{Q}$. This can be reduced to realizing the "Tits group" $\mathcal{T}$ of SL_n as a Galois group. $\mathcal{T}$ can be identified with the group of $n \times n$ signed permutation matrices with determinant one, which is an index two subgroup of the group of $n \times n$ signed permutation matrices. The latter is isomorphic to the Weyl group of type B_n, hence known to be a Galois group over $\mathbb{Q}$. As the N-T sequence splits if and only if n is odd, by Theorem 2.5, $\mathcal{T}$ can be realized over $\mathbb{Q}$ for n odd. When n is even, this problem is open except for small n, as far as the author knows.

Let $\mathfrak{g} = \mathfrak{sl}_n(k)$.

Proposition 2.9. *For l sufficiently large and $n \geq 4$, $\bar{\rho}(\mathfrak{g})$ decomposes into irreducible $\Gamma_{\mathbb{Q}}$-modules $\bar{\rho}(\mathfrak{t})$ and $\bar{\rho}(\mathfrak{g}_\Phi)$.*

Proof. This follows from Corollary 2.2 and Remark 2.3. $\square$

There remains the case when G is SL_2 or SL_3. For SL_3, see Section 2B. For SL_2, see Section 4B3.

2A4. Sp_{2n}. Let $G = \mathrm{Sp}_{2n}(k)$, then $\mathcal{W}$ is isomorphic to a semidirect product of S_n and $D := (\mathbb{Z}/2\mathbb{Z})^n$. We fix a maximal split torus T in $\mathrm{Sp}_{2n}(k)$. The N-T sequence does not split by [Adams and He 2017].

Lemma 2.10. *Consider the N-T sequence for Sp_{2n}. The group $S_n \subset \mathcal{W}$ has a section to N, whereas $D \subset \mathcal{W}$ does not have a section to N but there is a subgroup $\tilde{D}$ of N such that $\pi(\tilde{D}) = D$ and $\tilde{D} \cong (\mathbb{Z}/4\mathbb{Z})^n$. Moreover, as subgroups of N, S_n normalizes $\tilde{D}$ and $S_n \cap \tilde{D} = \{1\}$.*

We let $\mathcal{W}_1$ be the subgroup of N generated by S_n and $\tilde{D}$.

Proof. Let $V = k^{2n}$ be the $2n$-dimensional vector space over k endowed with a nondegenerate alternating form $(\cdot\,,\cdot)$. We may choose a basis

$$e_1, \ldots, e_n, e_1', \ldots, e_n'$$

of V such that $(e_i, e_j') = 1$ if and only if $i = j$, and that $(e_i, e_j) = 0$, $(e_i', e_j') = 0$ for all i, j. The Weyl group of Sp_{2n} is isomorphic to the semidirect product of the group S_n, which acts by permuting $e_1, \ldots, e_n$, and the group $D := (\mathbb{Z}/2\mathbb{Z})^n$, which acts by $e_i \mapsto (\pm 1)_i e_i$. There is an inclusion $S_n \to \mathrm{Sp}_{2n}(k)$ (which is a section to $N \to \mathcal{W}$) given as follows: $\forall \sigma \in S_n$, σ permutes $e_1, \ldots, e_n$ and $e_1', \ldots, e_n'$ by permuting the indices, which defines an element in $\mathrm{Sp}_{2n}(k)$. There is no such inclusion for D. However, we can define a 2-group $\tilde{D}$ which embeds into $\mathrm{Sp}_{2n}(k)$ as follows: for $1 \leq i \leq n$, let d_i be an endomorphism of V such that $d_i(e_i) = -e_i'$, $d_i(e_i') = e_i$, and d_i fixes all other basis vectors. It is clear that $d_i \in \mathrm{Sp}_{2n}(k)$. Let

$\tilde{D}$ be the subgroup of $\mathrm{Sp}_{2n}(k)$ generated by d_i for $1 \le i \le n$. Then $\tilde{D} \cong (\mathbb{Z}/4\mathbb{Z})^n$. It is obvious that (as subgroups of $\mathrm{Sp}_{2n}(k)$) S_n normalizes $\tilde{D}$ and $S_n \cap \tilde{D} = \{1\}$. □

Therefore, by Theorem 2.4 and 2.5, $\mathcal{W}_1$ can be realized as a Galois group over $\mathbb{Q}$ such that the complex conjugation corresponds to an element away from the center of G. The group N is generated by $\mathcal{W}_1$ and T. We have $\mathcal{W}_1$ normalizes T and $\mathcal{W}_1 \cap T \cong (\mathbb{Z}/2\mathbb{Z})^n$. Let S be the (abstract) semidirect product of $\mathcal{W}_1$ and T. By Theorem 2.5, for l large enough, S can be realized as a Galois group over $\mathbb{Q}$ unramified at l. Composing the corresponding map $\Gamma_{\mathbb{Q}} \twoheadrightarrow S$ with the natural surjection $S \twoheadrightarrow N$, we obtain a surjection $\Gamma_{\mathbb{Q}} \twoheadrightarrow N$ that is unramified at l and for which the complex conjugation maps to an element outside the center of G. Define $\bar{\rho}$ to be the composite

$$\Gamma_{\mathbb{Q}} \twoheadrightarrow N \xrightarrow{i} \mathrm{Sp}_{2n}(k).$$

Let $\mathfrak{g} = \mathfrak{sp}_{2n}(k)$. The root system Φ of $\mathfrak{g}$ is not simply laced.

Proposition 2.11. *For l sufficiently large, $\bar{\rho}(\mathfrak{g})$ decomposes into irreducible $\Gamma_{\mathbb{Q}}$-modules $\bar{\rho}(\mathfrak{t})$, $\bar{\rho}(\mathfrak{g}_l)$ and $\bar{\rho}(\mathfrak{g}_s)$.*

Proof. This follows from Corollary 2.2. □

2A5. Spin_{2n} *and* Spin_{2n+1}. For spin groups, the N-T sequence does not split by [Adams and He 2017]. For $G = \mathrm{Spin}_{2n}$, $\mathcal{W}$ is isomorphic to a semidirect product of S_n and $D := (\mathbb{Z}/2\mathbb{Z})^{n-1}$ and we let $\mathcal{W}'$ be the subgroup generated by A_n and D. For $G = \mathrm{Spin}_{2n+1}$, $\mathcal{W}$ is isomorphic to a semidirect product of S_n and $D := (\mathbb{Z}/2\mathbb{Z})^n$ and we let $\mathcal{W}'$ be the subgroup generated by A_n and D. Similar to the symplectic case, we will show that $N' = \pi^{-1}(\mathcal{W}')$ is a Galois group over $\mathbb{Q}$.

Lemma 2.12. *Consider the N-T sequence for $G = \mathrm{Spin}_{2n}(k)$ or $\mathrm{Spin}_{2n+1}(k)$. The map $\pi^{-1}(A_n) \to A_n$ admits a section, and there is a nilpotent subgroup $\tilde{D}$ of N such that $\pi(\tilde{D}) = D$. Moreover, as subgroups of N, A_n normalizes $\tilde{D}$ and we let $\mathcal{W}_1$ be their product.*

Proof. Let $\underline{G} := \mathrm{SO}_{2n}(k)$ or $\mathrm{SO}_{2n+1}(k)$. We have the standard homomorphism $i : \mathrm{GL}_n(k) \to \underline{G}$, which restricts to a homomorphism $S_n \to \underline{G}$. Let $\widetilde{\mathrm{GL}}_n(k)$ be the pullback of i along the covering map $G \to \underline{G}$. It is a two-fold central extension of $\mathrm{GL}_n(k)$, which can be identified with the group of pairs (g, z) with $g \in \mathrm{GL}_n(k)$, $z \in k^\times$, such that $\det g = z^2$, where the multiplication is defined by $(g_1, z_1) \cdot (g_2, z_2) = (g_1 g_2, z_1 z_2)$. A subgroup H of $\mathrm{GL}_n(k)$ has a section in $\widetilde{\mathrm{GL}}_n(k)$ if and only if the restriction of det to H is the square of a character of H. In particular, taking $H = A_n$, we see that A_n has a section to $\widetilde{\mathrm{GL}}_n(k)$. It follows that $\pi^{-1}(A_n) \to A_n$ admits a section.

The map $\underline{\pi}^{-1}(D) \to D$ has a section, where $\underline{\pi}$ is the natural map $N_{\underline{G}}(\underline{T}) \to \mathcal{W}$; this follows from [Adams and He 2017, Theorem 4.16], or can be seen directly from an elementary matrix calculation. Let $\tilde{D} \subset G$ be the preimage of $D \subset \underline{G}$ under the covering map $G \to \underline{G}$. As D is abelian, $[\tilde{D}, \tilde{D}] = Z(G) \cong \mu_2$. In particular, $\tilde{D}$ is nilpotent.

Finally, because A_n normalizes D in $\underline{G}$, A_n normalizes $\tilde{D}$ in G. □

Therefore, by Theorem 2.4 and Shafarevich's theorem (the group $\tilde{D}$ is nilpotent, see [Neukirch et al. 2000, IX, Section 6] for Shafarevich's theorem), $\mathcal{W}_1$ can be realized as a Galois group over $\mathbb{Q}$ such that the complex conjugation corresponds to an element outside the center of G. Let N' be the subgroup of G generated by $\mathcal{W}_1$ and T ($\mathcal{W}_1$ normalizes T). Let S be the (abstract) semidirect product of $\mathcal{W}_1$ and T. By Theorem 2.5, for l large enough, S can be realized as a Galois group over $\mathbb{Q}$ unramified at l. Composing the corresponding map $\Gamma_{\mathbb{Q}} \twoheadrightarrow S$ with the natural quotient map $S \twoheadrightarrow N'$, we obtain a surjection

$$\Gamma_{\mathbb{Q}} \twoheadrightarrow N'$$

that is unramified at l and for which the complex conjugation maps to an element outside the center of G (since the complex conjugation corresponds to an element outside the center of G in the realization of $\mathcal{W}_1$ as a Galois group over $\mathbb{Q}$). For $G = \mathrm{Spin}_{2n}$ or Spin_{2n+1}, define $\bar{\rho}$ to be the composite

$$\Gamma_{\mathbb{Q}} \twoheadrightarrow N' \xrightarrow{i} G(k).$$

Let $\mathfrak{g}$ be the Lie algebra of $G(k)$. The corresponding root system Φ is simply laced if $G = \mathrm{Spin}_{2n}$ and is not if $G = \mathrm{Spin}_{2n+1}$.

Proposition 2.13. *For l sufficiently large and $n \geq 4$, $\bar{\rho}(\mathfrak{g})$ decomposes into irreducible $\Gamma_{\mathbb{Q}}$-modules $\bar{\rho}(\mathfrak{t})$ and $\bar{\rho}(\mathfrak{g}_{\Phi})$ when $G = \mathrm{Spin}_{2n}$; and it decomposes into irreducible $\Gamma_{\mathbb{Q}}$-modules $\bar{\rho}(\mathfrak{t})$, $\bar{\rho}(\mathfrak{g}_l)$ and $\bar{\rho}(\mathfrak{g}_s)$ when $G = \mathrm{Spin}_{2n+1}$.*

Proof. Note that the action of $\mathcal{W}'$ on Φ is transitive if and only if $n \geq 4$. Then the proposition follows from Corollary 2.2 and Remark 2.3. □

Remark 2.14. There remains the case when G is one of Spin_4, Spin_5, Spin_6, Spin_7. But $\mathrm{Spin}_4(\bar{\mathbb{Q}}_l) \cong \mathrm{SL}_2(\bar{\mathbb{Q}}_l) \times \mathrm{SL}_2(\bar{\mathbb{Q}}_l)$, $\mathrm{Spin}_5(\bar{\mathbb{Q}}_l) \cong \mathrm{Sp}_4(\bar{\mathbb{Q}}_l)$, $\mathrm{Spin}_6(\bar{\mathbb{Q}}_l) \cong \mathrm{SL}_4(\bar{\mathbb{Q}}_l)$ which are included in other cases. For Spin_7, the half sum of coroots $\rho^\vee = 3\alpha_1^\vee + 5\alpha_2^\vee + 3\alpha_3^\vee$ has integer coefficients, so the principal GL_2 map is well defined, see Section 2B.

2A6. E_7^{sc}. Let $G = E_7^{\mathrm{sc}}(k)$. The Weyl group $\mathcal{W}$ is isomorphic to the direct product of $[\mathcal{W}, \mathcal{W}]$ and $\mathbb{Z}/2\mathbb{Z}$. By [Adams and He 2017], the N-T sequence does not split. We choose a subgroup $\mathcal{W}'$ of $\mathcal{W}$ which lifts to N as follows. Consider the extended Dynkin diagram of type E_7, there is a subroot system Φ' of Φ which is of type A_7. The alternating group A_8 is a subgroup of $S_8 \cong \mathcal{W}(A_7) \leq \mathcal{W} = \mathcal{W}(E_7)$.

Lemma 2.15. *The group $A_8 \leq \mathcal{W}$ lifts to N.*

Proof. This is because A_8 lifts to SL_8. □

Lemma 2.16. *The action of A_8 on Φ has an orbit of size 56 and an orbit of size 70.*

Proof. We first consider the action of $S_8 \cong \mathcal{W}(A_7)$ on Φ. By Lemma 2.1, S_8 acts transitively on Φ', which has 56 roots. A straightforward calculation using Plate E_7 in [Bourbaki 1968] shows that for some $\alpha \in \Phi - \Phi'$, $S_8 \cdot \alpha$ has exactly 70 roots (in the extended Dynkin diagram, take α to be the simple root that is not in Φ', then we let the group generated by the simple reflections in Φ' act on α and count the number of roots in the orbit). So the stabilizer of α in S_8 is isomorphic to $S_4 \times S_4 \subset S_8$. Since

$56+70=126$ is the number of roots in Φ, the lemma is true for S_8. Now we consider the alternating group A_8. Lemma 2.6 implies that A_8 still acts transitively on Φ'. As $(S_4 \times S_4) \cap A_8$ (the stabilizer of α in A_8) has order $\frac{1}{2}|S_4 \times S_4| = 288$, the orbit $A_8 \cdot \alpha$ has exactly $|A_8|/288 = 70$ roots. □

It is clear that A_8, considered as a subgroup of N, normalizes T and $A_8 \cap T = \{1\}$. Let N' be the subgroup of $G = E_7^{\mathrm{sc}}(k)$ generated by A_8 and T. By Theorems 2.4 and 2.5, for l large enough, we can find a continuous surjection

$$\Gamma_{\mathbb{Q}} \twoheadrightarrow N'$$

that is unramified at l and for which the complex conjugation maps to an element away from the center of G. Define $\bar{\rho}$ to be the composite

$$\Gamma_{\mathbb{Q}} \twoheadrightarrow N' \xrightarrow{i} E_7^{\mathrm{sc}}(k).$$

Let $\mathfrak{g}_a$ and $\mathfrak{g}_b$ be the direct sums of the root spaces corresponding to the orbit of size 56 and size 70, respectively, in Lemma 2.16.

Proposition 2.17. *For l sufficiently large, $\bar{\rho}(\mathfrak{g})$ decomposes into irreducible $\Gamma_{\mathbb{Q}}$-modules $\bar{\rho}(\mathfrak{t})$, $\bar{\rho}(\mathfrak{g}_a)$ and $\bar{\rho}(\mathfrak{g}_b)$.*

Proof. The proof is very similar to the proof of Corollary 2.2. □

2B. *The principal* GL$_2$ *construction.* We record some facts on the principal SL_2 and GL_2. For more details, see [Serre 1996, Section 1]. Let G/k be a simple algebraic group with a Borel B containing a split maximal torus T with unipotent radical U. Let $\Phi = \Phi(G, T)$ be the root system of G with the set of simple roots Δ corresponding to B. We fix a pinning $\{x_\alpha : \mathbb{G}_a \to U_\alpha\}$ where U_α is the root subgroup in B corresponding to α. Let $X_\alpha = dx_\alpha(1)$ for all $\alpha \in \Delta$ and let

$$X = \sum_{\alpha \in \Delta} X_\alpha,$$

which can be extended to an $\mathfrak{sl}_2$-triple (X, H, Y) where

$$H = \sum_{\alpha > 0} H_\alpha$$

with H_α the coroot vector corresponding to α.

When p is large enough relative to G, there is an exponential map

$$\exp : \mathrm{Lie}(U) \to U$$

which is an isomorphism.

A *principal* SL_2 *homomorphism* is a homomorphism

$$\varphi : \mathrm{SL}_2 \to G$$

such that

$$\varphi\left(\begin{pmatrix}1 & t\\ 0 & 1\end{pmatrix}\right)=\exp(tX),\quad \varphi\left(\begin{pmatrix}t & 0\\ 0 & t^{-1}\end{pmatrix}\right)=2\rho^\vee(t)$$

where $\rho^\vee$ is the half-sum of the coroots. $\rho^\vee$ is always defined when G is adjoint. Suppose that $\rho^\vee : \mathbb{G}_m \to G^{\mathrm{ad}}$ lifts to G and we fix a lift which is again denoted $\rho^\vee$. A *principal* GL_2 *homomorphism* is a homomorphism

$$\varphi : \mathrm{GL}_2 \to G$$

that extends a principal SL_2 such that such that

$$\varphi\left(\begin{pmatrix}1 & t\\ 0 & 1\end{pmatrix}\right)=\exp(tX),\quad \varphi\left(\begin{pmatrix}t & 0\\ 0 & 1\end{pmatrix}\right)=\rho^\vee(t).$$

By definition, a principal GL_2 factors through PGL_2.

By examining the list in [Bourbaki 1968], we get:

Lemma 2.18. *For G a simple algebraic group, $\rho^\vee : \mathbb{G}_m \to G^{\mathrm{ad}}$ lifts to G^{sc} if and only if G is one of the following types*: A_{2n}, B_{4n}, B_{4n+3}, D_{4n}, D_{4n+1}, E_6, E_8, F_4, G_2.

The operator $\mathrm{ad}(H)$ preserves $\mathfrak{g}^X$ (the centralizer of X in $\mathfrak{g}$) and

$$\mathfrak{g}^X=\sum_{m>0} V_{2m},$$

where V_{2m} is the eigenspace of H corresponding to the eigenvalue $2m$. The following proposition is due to Kostant [1959].

Proposition 2.19. *The dimension of $\mathfrak{g}^X$ is equal to the rank of $\mathfrak{g}$. V_{2m} is nonzero if and only if m is an exponent of $\mathfrak{g}$. Letting GL_2 act on $\mathfrak{g}$ via φ, there is an isomorphism of GL_2-representations*

$$\mathfrak{g}\cong\bigoplus_{m>0}\mathrm{Sym}^{2m}(k^2)\otimes\det{}^{-m}\otimes V_{2m}.$$

Suppose that $\rho^\vee : \mathbb{G}_m \to G^{\mathrm{ad}}$ lifts to G. We take f to be as in Example 1.5. By [Ribet 1985, Theorem 2.1], the projective image of $\bar{r}_{f,\lambda}$ is either $\mathrm{PGL}_2(k)$ or $\mathrm{PSL}_2(k)$ for a subfield k of k_λ. We then define

$$\bar{\rho} : \Gamma_{\mathbb{Q}} \xrightarrow{\bar{r}_{f,\lambda}} \mathrm{GL}_2(k) \xrightarrow{\varphi} G(k).$$

This construction works for all exceptional groups but E_7^{sc} as $\rho^\vee : \mathbb{G}_m \to E_7^{\mathrm{ad}}$ does not lift to E_7^{sc}. We will only use this construction for G of type E_6, A_2 and B_3.

3. Ramakrishna's method and its variants

Given $\bar{\rho} : \Gamma_{\mathbb{Q}} \to G(k)$ defined in the previous section, we want to obtain an l-adic lift $\rho : \Gamma_{\mathbb{Q}} \to G(\mathcal{O})$ with $\mathcal{O}$ the ring of integers of a finite extension of $\mathbb{Q}_l$ whose residue field is k satisfying a given global deformation condition. Just as in [Patrikis 2016], we use Ramakrishna's method to annihilate the associated dual Selmer group. The new feature is a double use of Patrikis' extension of Ramakrishna's method when the original form fails to work, see Section 3B.

3A. *Ramakrishna's method.*

3A1. *Ramakrishna deformations.* We list the key points and results of Patrikis' extension of Ramakrishna's method. For proofs, see [Patrikis 2016, Section 4.2]. For an overview on the deformation theory of (G-valued) Galois representations, see [loc. cit., Section 3].

We begin by defining a type of local deformation condition called Ramakrishna's condition, which will be imposed at the auxiliary primes of ramification in Ramakrishna's global argument. Let F be a finite extension of $\mathbb{Q}_p$ for $p \neq l$, and let $\bar{\rho} : \Gamma_F \to G(k)$ be an *unramified* homomorphism such that $\bar{\rho}(\mathrm{Fr}_F)$ is a regular semisimple element. Let T be the connected component of the centralizer of $\bar{\rho}(\mathrm{Fr}_F)$; this is a maximal k-torus of G, but we can lift it to an $\mathcal{O}$-torus uniquely up to isomorphism, which we also denote by T, and then we can lift the embedding over k to an embedding over $\mathcal{O}$ which is unique up to $\hat{G}(\mathcal{O})$-conjugation. By passing to an étale extension of $\mathcal{O}$, we may assume that T is split.

The following definition is from [loc. cit.].

Definition 3.1. Let $\bar{\rho}$, T be as above. For $\alpha \in \Phi(G, T)$, $\bar{\rho}$ is said to be of Ramakrishna type α if

$$\alpha(\bar{\rho}(\mathrm{Fr}_F)) = \bar{\kappa}(\mathrm{Fr}_F).$$

Let $H_\alpha = T \cdot U_\alpha$ be the subgroup generated by T and the root subgroup U_α corresponding to α. Ramakrishna deformation is a functor

$$\mathrm{Lift}^{\mathrm{Ram}}_{\bar{\rho}} : \mathrm{CNL}_{\mathcal{O}} \to \mathbf{Sets}$$

such that for a complete local noetherian $\mathcal{O}$-algebra R, $\mathrm{Lift}^{\mathrm{Ram}}_{\bar{\rho}}(R)$ consists of all lifts

$$\rho : \Gamma_F \to G(R)$$

of $\bar{\rho}$ such that ρ is $\hat{G}(R)$-conjugate to a homomorphism $\Gamma_F \xrightarrow{\rho'} H_\alpha(R)$ with the resulting composite

$$\Gamma_F \xrightarrow{\rho'} H_\alpha(R) \xrightarrow{\mathrm{Ad}} \mathrm{GL}(\mathfrak{g}_\alpha \otimes R) = R^\times$$

equal to κ.

We shall call such a ρ to be of Ramakrishna type α as well. We denote by $\mathrm{Def}^{\mathrm{Ram}}_{\bar{\rho}}$ the corresponding deformation functor.

The following lemma is [loc. cit., Lemma 4.10].

Lemma 3.2. $\mathrm{Lift}^{\mathrm{Ram}}_{\bar{\rho}}$ *is well defined and smooth.*

Consider the subtorus $T_\alpha = \mathrm{Ker}(\alpha)^0$ of T, and denote by $\mathfrak{t}_\alpha$ its Lie algebra. There is a canonical decomposition $\mathfrak{t}_\alpha \oplus \mathfrak{l}_\alpha = \mathfrak{t}$ with $\mathfrak{l}_\alpha$ the one-dimensional torus generated by the coroot $\alpha^\vee$.

The next lemma [loc. cit., Lemma 4.11] is crucial in the global deformation theory.

Lemma 3.3. *Assume $\bar{\rho}$ is of Ramakrishna type α. Let $W = \mathfrak{t}_\alpha \oplus \mathfrak{g}_\alpha$, and let $W^\perp$ be the annihilator of W under the Killing form on $\mathfrak{g}$. Let $L_{\bar{\rho}}^{\mathrm{Ram}}$ and $L_{\bar{\rho}}^{\mathrm{Ram},\perp}$ be the tangent space of $\mathrm{Def}_{\bar{\rho}}^{\mathrm{Ram}}$ and the annihilator of $L_{\bar{\rho}}^{\mathrm{Ram}}$ under the local duality pairing, respectively. Then*:

(1) $L_{\bar{\rho}}^{\mathrm{Ram}} \cong H^1(\Gamma_F, \bar{\rho}(W))$.

(2) $\dim L_{\bar{\rho}}^{\mathrm{Ram}} = h^0(\Gamma_F, \bar{\rho}(\mathfrak{g}))$.

(3) $L_{\bar{\rho}}^{\mathrm{Ram},\perp} \cong H^1(\Gamma_F, \bar{\rho}(W^\perp)(1))$.

(4) *Let $L_{\bar{\rho}}^{\mathrm{Ram},\square}$ and $L_{\bar{\rho}}^{\mathrm{Ram},\perp,\square}$ be the preimages in $Z^1(\Gamma_F, \bar{\rho}(\mathfrak{g}))$ of $L_{\bar{\rho}}^{\mathrm{Ram}}$ and $L_{\bar{\rho}}^{\mathrm{Ram},\perp}$, respectively. Under the canonical decomposition*
$$\mathfrak{g} = \bigoplus_{\gamma} \mathfrak{g}_\gamma \oplus \mathfrak{t}_\alpha \oplus \mathfrak{l}_\alpha,$$
all cocycles in $L_{\bar{\rho}}^{\mathrm{Ram},\square}$ and $L_{\bar{\rho}}^{\mathrm{Ram},\perp,\square}$ have $\mathfrak{l}_\alpha$ and $\mathfrak{g}_{-\alpha}$ components equal to zero, respectively.

3A2. *The global argument.* In this section, we assume G is semisimple. Let $\bar{\rho} : \Gamma_{\mathbb{Q},S} \to G(k)$ be a continuous homomorphism for which $h^0(\Gamma_{\mathbb{Q}}, \bar{\rho}(\mathfrak{g})) = h^0(\Gamma_{\mathbb{Q}}, \bar{\rho}(\mathfrak{g})(1)) = 0$. In particular, the deformation functor is representable. The following theorem is proved in [Patrikis 2016, Proposition 5.2]. For a review on global deformation theory and systems of Selmer groups, see [Patrikis 2016, Sections 3.2–3.3].

Theorem 3.4. *Suppose that there is a global deformation condition $\mathcal{L} = \{L_v\}_{v\in S}$ consisting of* smooth *local deformation conditions for each place $v \in S$. Let $K = \mathbb{Q}(\bar{\rho}(\mathfrak{g}), \mu_l)$. We assume the following*:

(1)
$$\sum_{v\in S} (\dim L_v) \geq \sum_{v\in S} h^0(\Gamma_{\mathbb{Q}_v}, \bar{\rho}(\mathfrak{g})).$$

(2) $H^1(\mathrm{Gal}\,(K/\mathbb{Q}), \bar{\rho}(\mathfrak{g}))$ *and* $H^1(\mathrm{Gal}\,(K/\mathbb{Q}), \bar{\rho}(\mathfrak{g})(1))$ *vanish.*

(3) *Assume item* (2) *holds. For any pair of nonzero Selmer classes $\phi \in H^1_{\mathcal{L}^\perp}(\Gamma_{\mathbb{Q},S}, \bar{\rho}(\mathfrak{g})(1))$ and $\psi \in H^1_{\mathcal{L}}(\Gamma_{\mathbb{Q},S}, \bar{\rho}(\mathfrak{g}))$, we can restrict them to Γ_K where they become homomorphisms, which are nonzero by item* (2). *Letting K_ϕ/K and K_ψ/K be their fixed fields, we assume that K_ϕ and K_ψ are linearly disjoint over K.*

(4) *Consider any ϕ and ψ as in the hypothesis of item* (3). *There exists an element $\sigma \in \Gamma_{\mathbb{Q}}$ such that $\bar{\rho}(\sigma)$ is a regular semisimple element of G, the connected component of whose centralizer we denote T, and such that there exists a root $\alpha \in \Phi(G, T)$ satisfying*

   (a) $\alpha(\bar{\rho}(\sigma)) = \bar{\kappa}(\sigma)$,

   (b) *$k[\psi(\Gamma_K)]$ has an element with nonzero $\mathfrak{l}_\alpha$ component, and*

   (c) *$k[\phi(\Gamma_K)]$ has an element with nonzero $\mathfrak{g}_{-\alpha}$ component.*

Then there exists a finite set of primes Q disjoint from S, and a lift $\rho : \Gamma_{\mathbb{Q},S\cup Q} \to G(\mathcal{O})$ of $\bar{\rho}$ such that ρ is of type L_v at all $v \in S$ and of Ramakrishna type at all $v \in Q$.

Proof. We sketch the proof for the reader's convenience. By the arguments in [Taylor 2003, Lemma 1.1] (which carry without modification to other groups) it suffices to enlarge $\mathcal{L}$ to make the dual Selmer group

$H^1_{\mathcal{L}^\perp}(\Gamma_{\mathbb{Q},S}, \bar\rho(\mathfrak{g})(1))$ vanish. We may assume the dual Selmer group is nontrivial and take a nonzero class ϕ in it. Item (1) implies by Wiles' formula (Proposition 4.10) that $H^1_{\mathcal{L}}(\Gamma_{\mathbb{Q},S}, \bar\rho(\mathfrak{g}))$ is nontrivial. So we can take a nonzero class ψ in it. Item (3), (4) and Chebotarev's density theorem all together imply there exists infinitely many $w \notin S$ such that $\psi \mid w \notin L^{\mathrm{Ram}}_{\bar\rho\mid w}$ and $\phi \mid w \notin L^{\mathrm{Ram},\perp}_{\bar\rho\mid w}$. In particular, we have

$$\psi \notin H^1_{\mathcal{L}\cup L^{\mathrm{Ram}}_{\bar\rho\mid w}}(\Gamma_{\mathbb{Q},S\cup w}, \bar\rho(\mathfrak{g})), \tag{1}$$

and

$$\phi \notin H^1_{\mathcal{L}^\perp\cup L^{\mathrm{Ram},\perp}_{\bar\rho\mid w}}(\Gamma_{\mathbb{Q},S\cup w}, \bar\rho(\mathfrak{g})(1)). \tag{2}$$

If we can show

$$H^1_{\mathcal{L}^\perp\cup L^{\mathrm{Ram},\perp}_{\bar\rho\mid w}}(\Gamma_{\mathbb{Q},S\cup w}, \bar\rho(\mathfrak{g})(1)) \subset H^1_{\mathcal{L}^\perp}(\Gamma_{\mathbb{Q},S}, \bar\rho(\mathfrak{g})(1)), \tag{3}$$

then (2) will imply that (3) is a strict inclusion. The key point now is that if we let L^{unr}_w denote the unramified cohomology at w, then $L_w = L^{\mathrm{unr}}_w \cap L^{\mathrm{Ram}}_{\bar\rho\mid w}$ is codimension one in L^{unr}_w, which, together with a double invocation of Wiles' formula and (1), implies

$$H^1_{\mathcal{L}^\perp}(\Gamma_{\mathbb{Q},S}, \bar\rho(\mathfrak{g})(1)) = H^1_{\mathcal{L}^\perp\cup L^\perp_w}(\Gamma_{\mathbb{Q},S\cup w}, \bar\rho(\mathfrak{g})(1)),$$

from which (3) follows. A variation of this argument can be found in the proof of Proposition 3.13. Now for the new Selmer system, item 1 still holds (Lemma 3.3(2)). So we can apply the above argument finitely many times until the dual Selmer group of the enlarged Selmer system vanishes. □

3B. *A variant of the global argument.* In this section, we let G be a simple algebraic group and let $\bar\rho : \Gamma_{\mathbb{Q}} \to G(k)$ be as in Section 2A. Recall that $\bar\rho(\Gamma_{\mathbb{Q}})$ is a subgroup of $N_G(T)_k$. For $l = \mathrm{char}(k)$ large enough, $\bar\rho(\mathfrak{g})$ decomposes into the sum of $\bar\rho(\mathfrak{t})$ and another one or two summands depending on whether or not $\Phi(G,T)$ is simply laced; see Propositions 2.9, 2.11, 2.13 and 2.17. We fix a Selmer system $\mathcal{L}$.

Proposition 3.5. *Assume that l is large enough. Then items* (2) *and* (3) *in Theorem 3.4 are satisfied.*

Proof. For item (2), note that $|\mathrm{Gal}\,(K/\mathbb{Q})|$ divides $(l-1)|\bar\rho(\Gamma_{\mathbb{Q}})|$, which is prime to l by the construction of $\bar\rho$. Since the coefficients field k of H^1 has characteristic l, this implies the vanishing of H^1.

For item (3), since $\psi : \mathrm{Gal}\,(K_\psi/K) \cong \psi(\Gamma_K)$ and $\phi : \mathrm{Gal}\,(K_\phi/K) \cong \phi(\Gamma_K)$ are $\mathrm{Gal}\,(K/\mathbb{Q})$-equivariant isomorphisms, it is enough to check that the irreducible summands in $\mathfrak{g}$ and $\mathfrak{g}(1)$ are nonisomorphic. We check this case by case. If G is of type A_n or D_n, by the construction of $\bar\rho$, the alternating group A_{n+1} or A_n, respectively, may be identified with a subgroup of $\bar\rho(\Gamma_{\mathbb{Q}})$. We take an element $\sigma \in \Gamma_{\mathbb{Q}}$ such that $\bar\rho(\sigma) \in A_n$ has order 2. Since $\mathbb{Q}(\bar\rho)$ (the fixed field of $\bar\rho$) is unramified at l, $\mathbb{Q}(\bar\rho)$ and $\mathbb{Q}(\mu_l)$ are linearly disjoint over $\mathbb{Q}$, so we may modify σ if necessary to make $\bar\kappa(\sigma) \neq 1$. Consider the eigenvalues of σ on $\mathfrak{t}$ and $\mathfrak{t}(1)$ (here we recall that $\mathfrak{t}$ is the Lie algebra of the maximal split torus T of G in the construction of $\bar\rho$); the eigenvalues on $\mathfrak{t}$ are ± 1, whereas none of the eigenvalues on $\mathfrak{t}(1)$ can be 1 or -1. Thus $\mathfrak{t}$ is not isomorphic to $\mathfrak{t}(1)$ as Galois modules. On the other hand, since $T := T(k) \subset \bar\rho(\Gamma_{\mathbb{Q}})$, we can find $\tau \in \Gamma_{\mathbb{Q}}$ such that $\bar\rho(\tau)$ is a regular semisimple element for which $\alpha(\bar\rho(\tau)) = a$ for $\alpha \in \Delta$, where a is a generator of $(\mathbb{Z}/l\mathbb{Z})^\times$ and Δ is a fixed set of simple roots in Φ. Again since $\mathbb{Q}(\bar\rho)$ and $\mathbb{Q}(\mu_l)$ are linearly disjoint over $\mathbb{Q}$, we may modify τ if necessary to make $\bar{\kappa}(\tau)=a$. Consider the eigenvalues of τ on $\mathfrak{g}_\Phi$ and $\mathfrak{g}_\Phi(1)$, those on $\mathfrak{g}_\Phi$ are $\{\alpha(\bar{\rho}(\tau)):\alpha\in\Phi\}$, whereas those on $\mathfrak{g}_\Phi(1)$ are $\{a\cdot\alpha(\bar{\rho}(\tau)):\alpha\in\Phi\}$. Note that $\alpha(\bar{\rho}(\tau))=a^{\mathrm{ht}(\alpha)}$ and the order of a is $l-1$, it is clear that these two sets are different when l is large enough. Thus, $\mathfrak{g}_\Phi$ and $\mathfrak{g}_\Phi(1)$ are nonisomorphic as Galois modules.

If G is of type B_n, by the construction of $\bar{\rho}$, the alternating group A_n may be identified with a subgroup of $\bar{\rho}(\Gamma_\mathbb{Q})$. Using the argument in the previous paragraph, we see that $\mathfrak{t}$ is not isomorphic to $\mathfrak{t}(1)$ as Galois modules. On the other hand, we need to show that $\bar{\rho}(\mathfrak{g}_l)$, $\bar{\rho}(\mathfrak{g}_s)$, $\bar{\rho}(\mathfrak{g}_l)(1)$, $\bar{\rho}(\mathfrak{g}_s)(1)$ are pairwise nonisomorphic as Galois modules. Just like before we can find $\tau\in\Gamma_\mathbb{Q}$ such that $\bar{\rho}(\tau)$ is a regular semisimple element for which $\alpha(\bar{\rho}(\tau))=a$ for $\alpha\in\Delta$ and $\bar{\kappa}(\tau)=a$, where a is a generator of $(\mathbb{Z}/l\mathbb{Z})^\times$ and Δ is a fixed set of simple roots in Φ. Since l is large enough, the sets $\{\alpha(\bar{\rho}(\tau)):\alpha\in\Phi_l\}$, $\{\alpha(\bar{\rho}(\tau)):\alpha\in\Phi_s\}$, $\{a\cdot\alpha(\bar{\rho}(\tau)):\alpha\in\Phi_l\}$ and $\{a\cdot\alpha(\bar{\rho}(\tau)):\alpha\in\Phi_s\}$ must be distinct. So τ has different eigenvalues on $\bar{\rho}(\mathfrak{g}_l)$, $\bar{\rho}(\mathfrak{g}_s)$, $\bar{\rho}(\mathfrak{g}_l)(1)$, $\bar{\rho}(\mathfrak{g}_s)(1)$ and hence they are pairwise nonisomorphic Galois modules.

The demonstrations are the same for type C_n and E_7. □

Let M be a finite dimensional k-vector space with a continuous $\Gamma_\mathbb{Q}$-action. Define its *Tate dual* to be the space $M^\vee=\mathrm{Hom}(M,\mu_\infty)$ equipped with the following $\Gamma_\mathbb{Q}$-action:

$$(\sigma f)(m):=\sigma(f(\sigma^{-1}m)).$$

Proposition 3.6. *For any continuous homomorphism $\bar{\rho}:\Gamma_\mathbb{Q}\to G(k)$, $\bar{\rho}(\mathfrak{g})^\vee\cong\bar{\rho}(\mathfrak{g})(1)$. For $\bar{\rho}$ as in Section 2A, $\bar{\rho}(\mathfrak{t})^\vee\cong\bar{\rho}(\mathfrak{t})(1)$.*

Proof. As l is sufficiently large, the killing form is a nondegenerate G-invariant symmetric bilinear form on $\mathfrak{g}$, which identifies the contragredient representation $\mathfrak{g}^*$ with $\mathfrak{g}$, and hence identifies $\bar{\rho}(\mathfrak{g})^\vee$ with $\bar{\rho}(\mathfrak{g})(1)$ as Galois modules. If $\bar{\rho}$ is as in Section 2A, then the Galois action on $\mathfrak{t}$ factors through W. It is easy to see that the standard bilinear form on $\mathfrak{t}$ is nondegenerate and W-invariant. Just as above, we deduce that $\bar{\rho}(\mathfrak{t})^\vee\cong\bar{\rho}(\mathfrak{t})(1)$ as Galois modules. □

Definition 3.7. Let $\mathcal{L}=\{L_v\}_{v\in S}$ be the Selmer system corresponding to a global deformation condition for $\bar{\rho}$ that is unramified outside a finite set of places S, and let $\mathcal{L}^\perp=\{L_v^\perp\}_{v\in S}$ be the associated dual Selmer system. Define the *M-Selmer group* as follows:

$$H^1_\mathcal{L}(\Gamma_S,M)=\mathrm{Ker}\left(H^1(\Gamma_S,M)\to\bigoplus_{v\in S}H^1(\Gamma_v,M)/(L_v\cap H^1(\Gamma_v,M))\right),$$

and define the *M-dual Selmer group* as follows:

$$H^1_{\mathcal{L}^\perp}(\Gamma_S,M^\vee)=\mathrm{Ker}\left(H^1(\Gamma_S,M^\vee)\to\bigoplus_{v\in S}H^1(\Gamma_v,M^\vee)/(L_v^\perp\cap H^1(\Gamma_v,M^\vee))\right).$$

To apply Theorem 3.4, we need to make sure that items (1)–(4) in it are satisfied. By choosing an appropriate $\mathcal{L}$, we can make item (1) hold. Items (2) and (3) are satisfied by Proposition 3.5. It is tricky to deal with item (4): the images of ϕ and ψ, which are $\Gamma_\mathbb{Q}$-submodules of $\bar{\rho}(\mathfrak{g})$, must satisfy the group-theoretic properties in (b) and (c); if we can find an element σ as in item (4) such that *all* summands of $\bar{\rho}(\mathfrak{g})$ satisfies these properties, then item (4) will be satisfied. So achieving item (4) crucially depends on the group-theoretic properties of submodules of $\bar{\rho}(\mathfrak{g})$. We need to find a regular semisimple element Σ in $\bar{\rho}(\Gamma_{\mathbb{Q}})$, the connected component of whose centralizer we denote T', for which there exists a root $\alpha' \in \Phi' := \Phi(G, T')$ such that:

(1) $\alpha'(\Sigma) \in (\mathbb{Z}/l\mathbb{Z})^\times$,

(2) Every irreducible summand of $\bar{\rho}(\mathfrak{g})$ has an element with nonzero $\mathfrak{l}_{\alpha'}$ component.

(3) Every irreducible summand of $\bar{\rho}(\mathfrak{g})$ has an element with nonzero $\mathfrak{g}_{-\alpha'}$ component.

Unfortunately, for our residual representation, there is no such element which meets all three conditions. The rest of this section will show how item (4) of Theorem 3.4 can be met by controlling $\psi(\Gamma_K)$ for a given class ψ in the Selmer group.

If we take a regular semisimple element Σ in T (which is the fixed split maximal torus of $G(k)$ we use when constructing the residual representation), then there is no $\alpha \in \Phi$ fulfilling both (2) and (3). Instead, we look for $\Sigma \in N = N_G(T)$ for which $\pi(\Sigma)$ is a nontrivial element in W, where $\pi : N_G(T) \twoheadrightarrow \mathcal{W}$ is the canonical quotient map.

Lemma 3.8. *Assume the characteristic of k is large enough for G. Then for any $w \in \mathcal{W}$ that fixes a noncentral element of* $\mathrm{Lie}(T)$, *there exist a regular semisimple element $n \in N$ (regular with respect to G) such that $\pi(n) = w$. If G is of type A_1, A_2, B_2, C_2, then for any $w \in \mathcal{W}$, there exist a regular semisimple element $n \in N$ (regular with respect to G) such that $\pi(n) = w$.*

Proof. The second part follows from a straightforward calculation. We will prove the first part. We first make the following observation: if M is a Levi subgroup of G, then (by looking at the action of simple roots outside M on elements of $Z_G(M)^0$) for every M-regular semisimple element $t \in M$, there is a G-regular semisimple element of $tZ_G(M)^0$. If w fixes a noncentral element of $\mathrm{Lie}(T)$, then we take M to be $Z_G(\mathrm{Lie}(T)^w)$, which is a proper Levi subgroup of G since the characteristic of k is large enough for G. By induction on the semisimple rank, there is a M-regular semisimple element n' such that $\pi(n') = w$, so there is a G-regular semisimple element $n = n'z$ with $z \in Z_G(M)^0$, and hence $\pi(n) = \pi(n'z) = \pi(n')\pi(z) = \pi(n') = w$. □

Remark 3.9. The above lemma should be true without assuming w fixes a noncentral element of $\mathrm{Lie}(T)$, but the author does not know how to remove this assumption. For GL_n, one can show that (by matrix calculations) the property holds for all $w \in S_n$ as long as the characteristic of k is large enough.

Let t' be a regular semisimple element in G and $t \in T$ be an element that is conjugate to t'. Then t and t' determine a unique bijection between $\Phi = \Phi(G, T)$ and $\Phi' = \Phi(G, T')$ with $T' = Z_G(t')^0$: for any $\alpha \in \Phi$, define $\alpha' \in \Phi'$ such that

$$\alpha'(h) = \alpha(g^{-1}hg)$$

for any $h \in T'$, where g is an element in G such that $g^{-1}t'g = t$. Since t is regular semisimple, α' is independent of g.

Recall that for $\alpha \in \Phi$, s_α is the simple reflection in the Weyl group of Φ associated to α.

Lemma 3.10. *Assume the characteristic of k is not 2. For a long root $\alpha \in \Phi$, let Σ be a regular semisimple element in N such that $\pi(\Sigma) = s_\alpha$ (which exists by Lemma 3.8). We fix an element $t \in T$ that is conjugate to Σ and let $T' = Z_G(\Sigma)^0$. The elements Σ and t determine a bijection between Φ and Φ' as above.*

- *If Φ is of type A_n ($n \geq 1$) or D_n ($n \geq 4$), then the root $\alpha' \in \Phi'$ corresponding to α fulfills* (1) *and* (3). *For* (2), *$\mathfrak{g}_\Phi$ (recall that $\mathfrak{g}_\Phi := \sum_{\alpha \in \Phi} \mathfrak{g}_\alpha$) has an element with nonzero $\mathfrak{l}_{\alpha'}$ component, but $\mathfrak{t}$ does not.*
- *If Φ is of type B_n, C_n ($n \geq 2$) or E_7, then α' fulfills* (1) *and $\mathfrak{t}$ has a vector with nonzero $\mathfrak{g}_{-\alpha'}$ component.*

Moreover, $\mathfrak{t}' \cap \mathfrak{t} = \mathfrak{t}'_{\alpha'} \cap \mathfrak{t} = W \cap \mathfrak{t}$, where $W = \mathfrak{t}'_{\alpha'} \oplus \mathfrak{g}_{\alpha'}$.

Proof. It is clear that $\alpha'(\Sigma) = -1$, so (1) is satisfied. We need to show:

- The space $\mathfrak{l}_\alpha$ has nonzero $\mathfrak{g}_{-\alpha'}$ component.
- The space $\mathfrak{g}_\alpha$ has nonzero $\mathfrak{l}_{\alpha'}$ and $\mathfrak{g}_{-\alpha'}$ component.
- $\mathfrak{t}' \cap \mathfrak{t} = \mathfrak{t}'_{\alpha'} \cap \mathfrak{t} = W \cap \mathfrak{t}$.

This is essentially a GL_2-calculation. We may perform the calculation in the subalgebra of $\mathfrak{g}$ generated by $\mathfrak{g}_\alpha$, whose root lattice is isomorphic to $\{x_1e_1 + x_2e_2 \mid x_i \in \mathbb{Z}, x_1 + x_2 = 0\}$. We take $\alpha = e_1 - e_2$ with the corresponding root vector

$$X_\alpha := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

We have $\Sigma = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, $P = \left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right)$. Then $P^{-1}\Sigma P = \mathrm{diag}(1, -1)$. The first bullet follows from the identity

$$P^{-1} \cdot \mathrm{diag}(h_1, h_2) \cdot P = -\tfrac{1}{2} \begin{pmatrix} -h_1 - h_2 & -h_1 + h_2 \\ -h_1 + h_2 & -h_1 - h_2 \end{pmatrix}.$$

The second bullet follows from the identity

$$P^{-1} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} P = -\tfrac{1}{2} \begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix}.$$

To show the third bullet, note that elements in $\mathfrak{t}'$ are of the form $\left(\begin{smallmatrix} h & k \\ k & h \end{smallmatrix}\right)$, elements in $\mathfrak{t}'_{\alpha'}$ are of the form $\left(\begin{smallmatrix} h & 0 \\ 0 & h \end{smallmatrix}\right)$, and elements in $\mathfrak{g}_{\alpha'}$ are of the form $\left(\begin{smallmatrix} x & -x \\ x & -x \end{smallmatrix}\right)$. It follows that all three intersections in the third bullet are the one dimensional k-vector space spanned by $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$. □

Lemma 3.11. *Suppose the characteristic of k is not 2 or 3. Assume that Φ is of type B_n, C_n ($n \geq 2$) or E_7, so $\bar{\rho}(\mathfrak{g}) = \bar{\rho}(\mathfrak{t}) \oplus \bar{\rho}(\mathfrak{g}_l) \oplus \bar{\rho}(\mathfrak{g}_s)$ for B_n and C_n, and $\bar{\rho}(\mathfrak{g}) = \bar{\rho}(\mathfrak{t}) \oplus \bar{\rho}(\mathfrak{g}_a) \oplus \bar{\rho}(\mathfrak{g}_b)$ for E_7 (see Section 2A4-2A6).*

(1) *(Type B_n and C_n) For a pair of nonperpendicular $\beta, \gamma \in \Phi$ with β long and γ short, let Σ be a regular semisimple element in N such that $\pi(\Sigma) = s_\beta \cdot s_\gamma$ (which exists by Lemma 3.8). We fix an element $t \in T$ that is conjugate to Σ and let $T' = Z_G(\Sigma)^0$. The elements Σ and t determine a*

bijection between Φ *and* $\Phi' = \Phi(G, T')$. *Then for all long root* α' *in the span of* β' *and* γ' (*which is a subsystem of* Φ' *of type* C_2), (3) *is satisfied. For* (2), $\mathfrak{g}_l$ *and* $\mathfrak{g}_s$ *have elements with nonzero* $\mathfrak{l}_{\alpha'}$ *component, but* $\mathfrak{t}$ *does not. The element* $\alpha'(\Sigma)$ *has order* 4 *in* $\bar{k}^\times$, *hence* (1) *is satisfied only for* $l \equiv 1(4)$.

(2) (*Type* E_7) *For a pair of nonperpendicular* $\beta, \gamma \in \Phi$ *with* $\mathfrak{g}_\beta \subset \mathfrak{g}_a$ *and* $\mathfrak{g}_\gamma \subset \mathfrak{g}_b$, *let* Σ *be a regular semisimple element in* N *such that* $\pi(\Sigma) = s_\beta \cdot s_\gamma$. *We fix an element* $t \in T$ *that is conjugate to* Σ *and let* $T' = Z_G(\Sigma)^0$. *The elements* Σ *and* t *determine a bijection between* Φ *and* $\Phi' = \Phi(G, T')$. *Then for all roots* α' *in the span of* β' *and* γ' (*which is a subsystem of* Φ' *of type* A_2), (3) *is satisfied. For* (2), $\mathfrak{g}_a$ *and* $\mathfrak{g}_b$ *have elements with nonzero* $\mathfrak{l}_{\alpha'}$ *component, but* $\mathfrak{t}$ *does not. The elements* $\alpha'(\Sigma)$ *has order* 3 *in* $\bar{k}^\times$, *hence* (1) *is satisfied only for* $l \equiv 1(3)$.

Moreover, $W \cap \mathfrak{t} \subset \mathfrak{t}'$, *where* $W = \mathfrak{t}'_{\alpha'} \oplus \mathfrak{g}_{\alpha'}$.

Proof. We first prove (2). Let α' be any root in the span of β' and γ'. We need to show:

- $\mathfrak{l}_\beta \oplus \mathfrak{l}_\gamma$ has nonzero $\mathfrak{g}_{-\alpha'}$ component.
- $\mathfrak{g}_\beta$ and $\mathfrak{g}_\gamma$ have nonzero $\mathfrak{l}_{\alpha'}$ component and nonzero $\mathfrak{g}_{-\alpha'}$ component.
- $W \cap \mathfrak{t} \subset \mathfrak{t}'$.

We may perform the calculation in the subalgebra of $\mathfrak{g}$ generated by $\mathfrak{g}_\beta$ and $\mathfrak{g}_\gamma$, whose root lattice is isomorphic to $\{x_1e_1 + x_2e_2 + x_3e_3 \mid x_i \in \mathbb{Z}, x_1 + x_2 + x_3 = 0\}$. We take $\beta = e_1 - e_2$, $\gamma = e_2 - e_3$ with the corresponding root vectors

$$X_\beta := \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad X_\gamma := \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

We have

$$\Sigma = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Let r be a (fixed) primitive 3-rd root of unity in $\bar{k}$, and let

$$P = \begin{pmatrix} 1 & 1 & 1 \\ 1 & r^2 & r \\ 1 & r & r^2 \end{pmatrix},$$

then

$$P^{-1}\Sigma P = \mathrm{diag}(1, r, r^2),$$

which implies $\alpha'(\Sigma)$ has order 3 in $\bar{k}^\times$. We have $P^{-1}\,\mathrm{diag}(a, b, c)P$ is a nonzero scalar multiple of

$$\begin{pmatrix} ar + br + cr & ar + b + cr^2 & ar + br^2 + c \\ ar + br^2 + c & ar + br + cr & ar + b + cr^2 \\ ar + b + cr^2 & ar + br^2 + c & ar + br + cr \end{pmatrix},$$

from which the first bullet follows. On the other hand, we have $P^{-1}X_\beta P$ is a nonzero scalar multiple of

$$\begin{pmatrix} r & 1 & r^2 \\ r & 1 & r^2 \\ r & 1 & r^2 \end{pmatrix},$$

from which it follows that $\mathfrak{g}_\beta$ has nonzero $\mathfrak{l}_{\alpha'}$ component and nonzero $\mathfrak{g}_{-\alpha'}$ component for any α'. Similarly, $\mathfrak{g}_\gamma$ has nonzero $\mathfrak{l}_{\alpha'}$ component and nonzero $\mathfrak{g}_{-\alpha'}$ component for any α'. The second bullet follows. We now show the third bullet for $\alpha' = \beta'$, the calculation is similar for other roots. We have

$$P \begin{pmatrix} a & x & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{pmatrix} P^{-1}$$

is a nonzero constant multiple of

$$\begin{pmatrix} 2ar+br+xr & -a+b+xr^2 & -ar^2+br^2+x \\ -ar^2+br^2+xr & br+xr^2 & -a+b+x \\ -a+b+xr & -ar^2+br^2+xr^2 & 2ar+br+x \end{pmatrix}.$$

A simple calculation shows demanding the off-diagonal entries in the above matrix to be zero will force all of a, b, x to be zero. Thus $W \cap \mathfrak{t}$ is trivial in the subalgebra of $\mathfrak{g}$ generated by $\mathfrak{g}_\beta$ and $\mathfrak{g}_\gamma$. It follows that $W \cap \mathfrak{t}$ is contained in $\mathfrak{t}'$.

The proof of (1) is very similar to that of (2). The computation may be performed in the subalgebra of $\mathfrak{g}$ generated by $\mathfrak{g}_\beta$ and $\mathfrak{g}_\gamma$, which is of type C_2. Let α' be a long root in the span of β' and γ'. We will show the three bullets above are true. The roots are $\{\pm(e_1-e_2), \pm(e_1+e_2), \pm 2e_1, \pm 2e_2\}$. We fix an alternating form $x_1y_4 + x_2y_3 - x_3y_2 - x_4y_1$ on k^2 and let $\beta = e_1 - e_2$ and $\gamma = 2e_1$ with corresponding root vectors

$$X_\beta := \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad X_\gamma := \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

We have

$$\Sigma = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Let r be a (fixed) primitive eighth root of unity in $\bar{k}$, and let

$$P = \begin{pmatrix} 1 & 1 & 1 & 1 \\ r & r^3 & r^5 & r^7 \\ r^3 & r & r^7 & r^5 \\ r^2 & r^6 & r^2 & r^6 \end{pmatrix},$$

then

$$P^{-1}\Sigma P = \mathrm{diag}(r, r^3, r^5, r^7),$$

which implies $\alpha'(\Sigma)$ has order 4 in $\bar{k}^\times$. We have $P^{-1}\,\mathrm{diag}(a, b, -b, -a)P$ is a nonzero scalar multiple of

$$\begin{pmatrix} 0 & a-br^6 & 0 & a-br^2 \\ a-br^6 & 0 & a+br^2 & 0 \\ 0 & a+br^6 & 0 & a+br^2 \\ a-br^6 & 0 & a-br^2 & 0 \end{pmatrix}.$$

Let $\alpha' \in \Phi'$ be any long root, then α' satisfies the first bullet. On the other hand, $P^{-1}X_\beta P$ is a nonzero scalar multiple of

$$\begin{pmatrix} r-r^7 & 0 & r^5-r^7 & r^7-r^3 \\ 0 & r^3+r & 2r^5 & r^7+r \\ r+r^7 & 2r^3 & r^5+r^7 & 0 \\ r-r^5 & r^3-r & 0 & r^7-r \end{pmatrix},$$

and $P^{-1}X_\gamma P$ is a nonzero scalar multiple of

$$\begin{pmatrix} r^2 & r^6 & r^2 & r^6 \\ r^2 & r^6 & r^2 & r^6 \\ r^2 & r^6 & r^2 & r^6 \\ r^2 & r^6 & r^2 & r^6 \end{pmatrix},$$

from which the second bullet follows. We check the third bullet for α' corresponding to $2e_2$, the calculation is similar for other long roots in Φ'. We have

$$P\begin{pmatrix} a & 0 & 0 & 0 \\ 0 & 0 & x & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -a \end{pmatrix}P^{-1}$$

is a nonzero scalar multiple of

$$\begin{pmatrix} x & a(r^5-r^3)+xr^3 & a(r^5-r^3)-xr^5 & a(r^6-r^2)+xr^6 \\ a(r-r^7)+xr^3 & r^6x & 2ar^6-x & a(r^5-r^3)+xr \\ a(r^3-r^5)+xr & 2ar^2-x & -r^6x & a(r^3-r^5)+xr^7 \\ -2ar^6+xr^6 & a(r^3-r^5)+xr & a(r^7-r)-xr^3 & -x \end{pmatrix}.$$

It is easy to see that for the matrix to be diagonal, both of a and x have to be zero. Thus $W\cap\mathfrak{t}$ is trivial in the subalgebra of $\mathfrak{g}$ generated by $\mathfrak{g}_\beta$ and $\mathfrak{g}_\gamma$. It follows that $W\cap\mathfrak{t}$ is contained in $\mathfrak{t}'$. □

Let Σ, α' be as in Lemma 3.10, (1) for Φ of type A_n or D_n, and Lemma 3.11 for Φ of type B_n, C_n or E_7. We have $\alpha(\Sigma)\in(\mathbb{Z}/l\mathbb{Z})^\times$ for primes l satisfying an appropriate congruence condition if necessary. We need to modify the element Σ to make it land in $\bar{\rho}(\Gamma_\mathbb{Q})$. When Φ is of type A_n or D_n, $\pi(\bar{\rho}(\Gamma_\mathbb{Q}))=[\mathcal{W},\mathcal{W}]$ (see 2.1.3 and 2.1.5). For any $\alpha\in\Phi$, we write $\mathfrak{sl}_2^\alpha$ for the Lie subalgebra of $\mathfrak{g}$ generated by $\mathfrak{g}_\alpha$ and $\mathfrak{g}_{-\alpha}$. We replace s_α by $s_\alpha s_\beta$ for some root β orthogonal to α such that $[\mathfrak{sl}_2^\alpha, \mathfrak{sl}_2^\beta]$ is trivial (such β exists because

$n \geq 4$) and replace Σ with a regular semisimple element in $N_G(T)$ that maps to $s_\alpha s_\beta$ when modulo $T(k)$. Note that $s_\alpha s_\beta \in [\mathcal{W}, \mathcal{W}]$: The Weyl group $\mathcal{W}$ acts transitively on the irreducible root system Φ, so there exists $w \in \mathcal{W}$ such that $w\alpha = \beta$, and hence $ws_\alpha w^{-1} = s_\beta$; it follows that $s_\alpha s_\beta = [s_\alpha, w]$. We again denote this new element by Σ, which now lands in $\bar{\rho}(\Gamma_\mathbb{Q})$. Lemma 3.10(1) still holds. When Φ is of type B_n, we again have $\pi(\bar{\rho}(\Gamma_\mathbb{Q})) = [\mathcal{W}, \mathcal{W}]$ (see 2.1.5). As $s_\beta s_\gamma \in [\mathcal{W}, \mathcal{W}]$, $\Sigma \in \bar{\rho}(\Gamma_\mathbb{Q})$, so no modification is needed. When Φ is of type C_n, $\pi(\bar{\rho}(\Gamma_\mathbb{Q})) = \mathcal{W}$ (see 2.1.4), so we automatically have $\Sigma \in \bar{\rho}(\Gamma_\mathbb{Q})$. When Φ is of type E_7, the corresponding element Σ is in $\bar{\rho}(\Gamma_\mathbb{Q})$ as well (see 2.1.6). Since $\bar{\rho}$ is unramified at l and $\mathbb{Q}(\mu_l)$ is totally ramified at l, $\mathbb{Q}(\bar{\rho})$ and $\mathbb{Q}(\mu_l)$ are linearly disjoint over $\mathbb{Q}$. So there exists an element $\sigma \in \Gamma_\mathbb{Q}$ such that $\bar{\rho}(\sigma) = \Sigma$ and $\bar{\kappa}(\sigma) = \alpha'(\Sigma)$. It follows that $\alpha'(\bar{\rho}(\sigma)) = \bar{\kappa}(\sigma)$.

Lemma 3.12. *Suppose there is a Selmer system $\mathcal{L} = \{L_v\}_{v \in S}$ for which the $\mathfrak{t}$-Selmer group $H^1_\mathcal{L}(\Gamma_S, \bar{\rho}(\mathfrak{t}))$ is trivial. We take a pair of nonzero Selmer classes $\phi \in H^1_{\mathcal{L}^\perp}(\Gamma_{\mathbb{Q},S}, \bar{\rho}(\mathfrak{g})(1))$ and $\psi \in H^1_\mathcal{L}(\Gamma_{\mathbb{Q},S}, \bar{\rho}(\mathfrak{g}))$. Then item* (4) *of Theorem 3.4 is satisfied.*

Proof. We need to check Theorem 3.4(4)(b) and (c). First, Proposition 3.5 and the inflation-restriction sequence imply that $\psi(\Gamma_K)$ and $\phi(\Gamma_K)$ are nontrivial. By Lemmas 3.10 and 3.11, every irreducible summand of $\bar{\rho}(\mathfrak{g})$ has an element with nonzero $\mathfrak{g}_{-\alpha'}$ component. In particular, (c) holds. As $H^1_\mathcal{L}(\Gamma_S, \bar{\rho}(\mathfrak{t}))$ is trivial, $\psi(\Gamma_K) \nsubseteq \mathfrak{t}$, which implies that $k[\psi(\Gamma_K)]$ contains $\mathfrak{g}_\Phi$ when Φ is of type A_n or D_n, $k[\psi(\Gamma_K)]$ contains one of $\mathfrak{g}_l$ and $\mathfrak{g}_s$ when Φ is of type B_n or C_n, and $k[\psi(\Gamma_K)]$ contains one of $\mathfrak{g}_a$ and $\mathfrak{g}_b$ when Φ is of type E_7. It then follows from Lemmas 3.10 and 3.11 that $k[\psi(\Gamma_K)]$ has an element with nonzero $\mathfrak{l}_{\alpha'}$ component. So (b) holds as well. ☐

The next proposition achieves the vanishing assumption of the $\mathfrak{t}$-Selmer in Lemma 3.12 by using of a variant of the cohomological arguments in Ramakrishna's method.

Proposition 3.13. *Suppose that*

$$h^1_\mathcal{L}(\Gamma_S, \bar{\rho}(\mathfrak{t})) \leq h^1_{\mathcal{L}^\perp}(\Gamma_S, \bar{\rho}(\mathfrak{t})(1)).$$

Then there is a finite set of places Q disjoint from S and a Ramakrishna deformation condition for each $w \in Q$ with tangent space L_w^{Ram} such that

$$H^1_{\mathcal{L} \cup \{L_w^{\mathrm{Ram}}\}_{w \in Q}}(\Gamma_{S \cup Q}, \bar{\rho}(\mathfrak{t})) = 0.$$

We may assume that $H^1_\mathcal{L}(\Gamma_S, \bar{\rho}(\mathfrak{t}))$ is nontrivial, for otherwise we are done. The inequality in Proposition 3.13 then implies that $H^1_{\mathcal{L}^\perp}(\Gamma_S, \bar{\rho}(\mathfrak{t})(1))$ is nontrivial. Let $0 \neq \phi \in H^1_{\mathcal{L}^\perp}(\Gamma_S, \bar{\rho}(\mathfrak{t})(1))$.

Lemma 3.14. *There exists $\tau \in \Gamma_\mathbb{Q}$ with the following properties*:

(1) *$\bar{\rho}(\tau)$ is a regular semisimple element of $G(k)$, the connected component of whose centralizer we denote T'.*

(2) *There exists $\alpha' \in \Phi(G, T')$, such that $\alpha'(\bar{\rho}(\tau)) = \bar{\kappa}(\tau)$.*

(3) *$k[\phi(\Gamma_K)]$ has an element with nonzero $\mathfrak{g}_{-\alpha'}$-component.*

Proof. We have seen that the groups $H^1(\mathrm{Gal}(K/\mathbb{Q}), \bar{\rho}(\mathfrak{g}))$ and $H^1(\mathrm{Gal}(K/\mathbb{Q}), \bar{\rho}(\mathfrak{g})(1))$ are both trivial. In particular, the groups $H^1(\mathrm{Gal}(K/\mathbb{Q}), \bar{\rho}(\mathfrak{t}))$ and $H^1(\mathrm{Gal}(K/\mathbb{Q}), \bar{\rho}(\mathfrak{t})(1))$ are both trivial. The restriction-inflation sequence then implies that $\phi(\Gamma_K)$ is nontrivial. Now we let Σ, α' be as in Lemma 3.10(1) for Φ of type A_n or D_n, and Lemma 3.10(2) for Φ of type B_n, C_n or E_7. If necessary, we can modify Σ to make it land in $\bar{\rho}(\Gamma_{\mathbb{Q}})$, as explained in the paragraph preceding Lemma 3.12. We have $\alpha(\Sigma) \in (\mathbb{Z}/l\mathbb{Z})^\times$. Since $\bar{\rho}$ is unramified at l and $\mathbb{Q}(\mu_l)$ is totally ramified at l, $\mathbb{Q}(\bar{\rho})$ and $\mathbb{Q}(\mu_l)$ are linearly disjoint over $\mathbb{Q}$. So there exists an element $\tau \in \Gamma_{\mathbb{Q}}$ such that $\bar{\rho}(\tau) = \Sigma$ and $\bar{\kappa}(\tau) = \alpha'(\Sigma)$. It follows that $\alpha'(\bar{\rho}(\tau)) = \bar{\kappa}(\tau)$, proves (2). Statement (3) follows from Lemma 3.10. □

Corollary 3.15. *There exist infinitely many places $w \notin S$ such that $\bar{\rho}|_{\Gamma_w}$ is of Ramakrishna type α' and $\phi|_{\Gamma_w} \notin L_w^{\mathrm{Ram},\perp}$.*

Proof. This follows from Lemmas 3.3, 3.14 and Chebotarev's density theorem. See the proof of [Patrikis 2016, Lemma 5.3]. □

Proof of Proposition 3.13. Let w be chosen as in Corollary 3.15. We will show that

$$h^1_{\mathcal{L}^\perp \cup L_w^{\mathrm{Ram},\perp}}(\Gamma_{S\cup w}, \bar{\rho}(\mathfrak{t})(1)) < h^1_{\mathcal{L}^\perp}(\Gamma_S, \bar{\rho}(\mathfrak{t})(1)) \tag{4}$$

and

$$h^1_{\mathcal{L}\cup L_w^{\mathrm{Ram}}}(\Gamma_{S\cup w}, \bar{\rho}(\mathfrak{t})) - h^1_{\mathcal{L}^\perp \cup L_w^{\mathrm{Ram},\perp}}(\Gamma_{S\cup w}, \bar{\rho}(\mathfrak{t})(1)) = h^1_{\mathcal{L}}(\Gamma_S, \bar{\rho}(\mathfrak{t})) - h^1_{\mathcal{L}^\perp}(\Gamma_S, \bar{\rho}(\mathfrak{t})(1)) \tag{5}$$

which imply that

$$h^1_{\mathcal{L}\cup L_w^{\mathrm{Ram}}}(\Gamma_{S\cup w}, \bar{\rho}(\mathfrak{t})) < h^1_{\mathcal{L}}(\Gamma_S, \bar{\rho}(\mathfrak{t})),$$

from which Proposition 3.13 follows by induction.

We first show (5). By a double invocation of Wiles' formula (see Proposition 4.10), the difference between the two sides equals $\dim(L_w^{\mathrm{Ram}} \cap H^1(\Gamma_w, \bar{\rho}(\mathfrak{t}))) - h^0(\Gamma_w, \bar{\rho}(\mathfrak{t})) = h^1(\Gamma_w, \bar{\rho}(W \cap \mathfrak{t})) - h^0(\Gamma_w, \bar{\rho}(\mathfrak{t}))$. As $H^0(\Gamma_w, \bar{\rho}(\mathfrak{g})) = \mathfrak{t}'$, we have $H^0(\Gamma_w, \bar{\rho}(\mathfrak{t})) = \mathfrak{t} \cap \mathfrak{t}'$; on the other hand, the action of $\bar{\rho}(\Gamma_w)$ on $W \cap \mathfrak{t}$ is a sum of the trivial representation and the cyclotomic character. By an elementary calculation in Galois cohomology, $h^1(\Gamma_w, k) = h^1(\Gamma_w, k(1)) = 1$. It follows that $h^1(\Gamma_w, \bar{\rho}(W \cap \mathfrak{t})) = \dim W \cap \mathfrak{t}$ and so $h^1(\Gamma_w, \bar{\rho}(W \cap \mathfrak{t})) - h^0(\Gamma_w, \bar{\rho}(\mathfrak{t})) = \dim(W \cap \mathfrak{t}) - \dim(\mathfrak{t}' \cap \mathfrak{t})$, which is zero by Lemma 3.10.

It remains to prove (4). Let $L_w = L_w^{\mathrm{unr}} \cap L_w^{\mathrm{Ram}}$, so $L_w^\perp = L_w^{\mathrm{unr},\perp} + L_w^{\mathrm{Ram},\perp}$. We have the following obvious inclusions

$$H^1_{\mathcal{L}\cup L_w}(\Gamma_{S\cup w}, \bar{\rho}(\mathfrak{t})) \subset H^1_{\mathcal{L}\cup L_w^{\mathrm{Ram}}}(\Gamma_{S\cup w}, \bar{\rho}(\mathfrak{t})), \tag{6}$$

$$H^1_{\mathcal{L}^\perp\cup L_w^{\mathrm{Ram},\perp}}(\Gamma_{S\cup w}, \bar{\rho}(\mathfrak{t})(1)) \subset H^1_{\mathcal{L}^\perp\cup L_w^\perp}(\Gamma_{S\cup w}, \bar{\rho}(\mathfrak{t})(1)), \tag{7}$$

$$H^1_{\mathcal{L}\cup L_w}(\Gamma_{S\cup w}, \bar{\rho}(\mathfrak{t})) \subset H^1_{\mathcal{L}\cup L_w^{\mathrm{unr}}}(\Gamma_{S\cup w}, \bar{\rho}(\mathfrak{t})) = H^1_{\mathcal{L}}(\Gamma_S, \bar{\rho}(\mathfrak{t})), \tag{8}$$

$$H^1_{\mathcal{L}^\perp}(\Gamma_S, \bar{\rho}(\mathfrak{t})(1)) = H^1_{\mathcal{L}^\perp\cup L_w^{\mathrm{unr},\perp}}(\Gamma_{S\cup w}, \bar{\rho}(\mathfrak{t})(1)) \subset H^1_{\mathcal{L}^\perp\cup L_w^\perp}(\Gamma_{S\cup w}, \bar{\rho}(\mathfrak{t})(1)). \tag{9}$$

As $\phi|_{\Gamma_w} \notin L_w^{\mathrm{Ram},\perp}$, (7) is a strict inclusion. We claim that (9) is an isomorphism, which will imply (4). To prove our claim, we consider (8) first. There is an exact sequence

$$0 \to H^1_{\mathcal{L}\cup L_w}(\Gamma_{S\cup w}, \bar{\rho}(\mathfrak{t})) \to H^1_{\mathcal{L}}(\Gamma_S, \bar{\rho}(\mathfrak{t})) \to (L_w^{\mathrm{unr}} \cap H^1(\Gamma_w, \bar{\rho}(\mathfrak{t})))/(L_w \cap H^1(\Gamma_w, \bar{\rho}(\mathfrak{t}))).$$

As

$$L_w^{\mathrm{unr}} = H^1(\Gamma_w/I_w, \bar{\rho}(\mathfrak{g})) \xrightarrow{f\mapsto f(\mathrm{Fr}_w)} \mathfrak{g}/(\bar{\rho}(\mathrm{Fr}_w)-1)\mathfrak{g} \cong \mathfrak{t}',$$

the top of its last term is isomorphic to $H^1(\Gamma_w, \bar{\rho}(\mathfrak{t}'\cap\mathfrak{t}))$, which has dimension $\dim(\mathfrak{t}'\cap\mathfrak{t})$; the bottom of its last term is isomorphic to $H^1(\Gamma_w, \mathfrak{t}'_\alpha\cap\mathfrak{t})$, which has dimension $\dim(\mathfrak{t}'_\alpha\cap\mathfrak{t})$. By Lemma 3.10, these dimensions are equal. So the last term is zero, and hence (8) is an isomorphism. A double invocation of Wiles' formula (Proposition 4.10) shows $h^1_{\mathcal{L}^\perp\cup L_w^\perp}(\Gamma_{S\cup w}, \bar{\rho}(\mathfrak{t})(1)) - h^1_{\mathcal{L}^\perp}(\Gamma_S, \bar{\rho}(\mathfrak{t})(1))$ equals

$$h^1_{\mathcal{L}\cup L_w}(\Gamma_{S\cup w}, \bar{\rho}(\mathfrak{t})) - h^1_{\mathcal{L}}(\Gamma_S, \bar{\rho}(\mathfrak{t})) + h^0(\Gamma_w, \bar{\rho}(\mathfrak{t})) - \dim(L_w \cap H^1(\Gamma_w, \bar{\rho}(\mathfrak{t}))).$$

Because (8) is an isomorphism and $\dim(L_w\cap H^1(\Gamma_w, \bar{\rho}(\mathfrak{t}))) = h^1(\Gamma_w, \bar{\rho}(W\cap\mathfrak{t}'\cap\mathfrak{t})) = \dim(W\cap\mathfrak{t}'\cap\mathfrak{t}) = \dim(\mathfrak{t}'\cap\mathfrak{t}) = h^0(\Gamma_w, \bar{\rho}(\mathfrak{t}))$ (Lemma 3.10), the right hand side of the above identity is zero. Therefore, (9) is an isomorphism, which completes the proof of the proposition. □

Theorem 3.16. *Let $\mathcal{L} = \{L_v\}_{v\in S}$ be a family of smooth local deformation conditions for $\bar{\rho}$ (the residual representation defined in Section 2A) unramified outside a finite set of places S containing the real place and all places where $\bar{\rho}$ is ramified. Suppose that*

$$\sum_{v\in S} \dim L_v \geq \sum_{v\in S} h^0(\Gamma_v, \bar{\rho}(\mathfrak{g})) \quad \textit{and} \quad \sum_{v\in S} \dim(L_v \cap H^1(\Gamma_v, \bar{\rho}(\mathfrak{t}))) \leq \sum_{v\in S} h^0(\Gamma_v, \bar{\rho}(\mathfrak{t})).$$

Assume l is large enough; in addition, if Φ is of type E_7, assume $l \equiv 1(3)$, and if Φ is doubly laced, assume $l \equiv 1(4)$.

Then there is a finite set of places Q disjoint from S and a continuous lift

$$\rho : \Gamma_{S\cup Q} \to G(\mathcal{O})$$

of $\bar{\rho}$ such that ρ is of type L_v for $v \in S$ and of Ramakrishna type for $v \in Q$.

Proof. The second inequality and Wiles' formula (Proposition 4.10) imply that

$$h^1_{\mathcal{L}}(\Gamma_S, \bar{\rho}(\mathfrak{t})) \leq h^1_{\mathcal{L}^\perp}(\Gamma_S, \bar{\rho}(\mathfrak{t})(1)).$$

By Proposition 3.13, we can enlarge $\mathcal{L}$ by adding finitely many Ramakrishna deformation conditions to get a new Selmer system $\mathcal{L}' = \{L_v\}_{v\in S'}$ with $S' \supset S$ such that $H^1_{\mathcal{L}'}(\Gamma_{S'}, \bar{\rho}(\mathfrak{t})) = \{0\}$. By Lemma 3.3(2), replacing $\mathcal{L}$ by $\mathcal{L}'$ preserves the first inequality.

We choose $\sigma \in \Gamma_{\mathbb{Q}}$ and $\alpha' \in \Phi(G, T')$ as in the paragraph preceding Lemma 3.12 (this is where the congruence conditions for Φ of type B_n, C_n, E_7 come in). Chebotarev's density theorem implies that there are infinitely many places $w \notin S'$ such that $\bar{\rho}|_w$ is of Ramakrishna type α'. We have for such a prime w,

$$H^1_{\mathcal{L}'\cup L_w^{\mathrm{Ram}}}(\Gamma_{S'\cup w}, \bar{\rho}(\mathfrak{t})) = 0.$$

In other words, adding a Ramakrishna local deformation conditions does not make the $\mathfrak{t}$-Selmer group jump back to a nontrivial group. Indeed, $L_w^{\mathrm{Ram}} \cap H^1(\Gamma_w, \bar{\rho}(\mathfrak{t})) = H^1(\Gamma_w, \bar{\rho}(W \cap \mathfrak{t})) \subset H^1(\Gamma_w, \bar{\rho}(\mathfrak{t}' \cap \mathfrak{t})) = H^1(\Gamma_w/I_w, \bar{\rho}(\mathfrak{t}))$, where the middle inclusion follows from Lemma 3.10 and 3.11. So

$$H^1_{\mathcal{L}' \cup L_w^{\mathrm{Ram}}}(\Gamma_{S' \cup w}, \bar{\rho}(\mathfrak{t})) \subset H^1_{\mathcal{L}'}(\Gamma_{S'}, \bar{\rho}(\mathfrak{t})) = \{0\}.$$

Let us check the assumptions of Theorem 3.4. By Proposition 4.10 and the first inequality in the assumption, item (1) holds. Item (2) and (3) are satisfied by Proposition 3.5. As $H^1_{\mathcal{L}'}(\Gamma_{S'}, \bar{\rho}(\mathfrak{t})) = \{0\}$, Item (4) is satisfied by Lemma 3.12. Therefore, by the proof of Theorem 3.4, there is a strict inclusion

$$H^1_{\mathcal{L}'^{\perp} \cup L_w^{\mathrm{Ram},\perp}}(\Gamma_{\mathbb{Q},S' \cup w}, \bar{\rho}(\mathfrak{g})(1)) \subset H^1_{\mathcal{L}'^{\perp}}(\Gamma_{\mathbb{Q},S'}, \bar{\rho}(\mathfrak{g})(1)).$$

As the $\mathfrak{t}$-Selmer group is still trivial for the enlarged Selmer system, item (4) remains valid by Lemma 3.12. So we can find a prime $w' \notin S' \cup w$ and enlarge the Selmer system $\mathcal{L}' \cup L_w^{\mathrm{Ram}}$ in the same way so that the dual Selmer group shrinks even further. Applying this argument finitely many times, we can kill the dual Selmer group. Therefore, by the first two lines of the proof of Theorem 3.4, we obtain desired l-adic lifts. □

3C. *Deforming principal* GL_2. We use the notation in Section 2B. Recall that $\bar{\rho}$ is the composite $\Gamma_{\mathbb{Q}} \to \mathrm{GL}_2(k) \xrightarrow{\varphi} G(k)$ where the first map is constructed from modular forms and the second map is the principal GL_2-map.

Patrikis has shown that all simple algebraic groups of exceptional types are geometric monodromy groups for $\Gamma_{\mathbb{Q}}$ except for E_6^{ad}, E_6^{sc}, E_7^{sc} [Patrikis 2016]. In this section, we follow Patrikis' work and use the principal GL_2 to construct full-image Galois representations into E_6^{ad}, E_6^{sc}, SL_3, Spin_7.

The proof of the following theorem is identical to that of [Patrikis 2016, Theorem 7.4].

Theorem 3.17. *Let $\mathcal{L} = \{L_v\}_{v \in S}$ be a family of smooth local deformation conditions for $\bar{\rho}$ (the residual representation defined in Section 2B) unramified outside a finite set of places S containing the real place and all places where $\bar{\rho}$ is ramified. Suppose that*

$$\sum_{v \in S} \dim L_v \geq \sum_{v \in S} h^0(\Gamma_v, \bar{\rho}(\mathfrak{g})).$$

Assume l is large enough.

Then there is a finite set of places Q disjoint from S and a continuous lift

$$\rho : \Gamma_{S \cup Q} \to G(\mathcal{O})$$

of $\bar{\rho}$ such that ρ is of type L_v for $v \in S$ and of Ramakrishna type for $v \in Q$.

In [Patrikis 2016, Lemma 7.6], the following fact is verified using Magma.

Lemma 3.18. *Assume l is large enough for $\mathfrak{g}$. For $\mathfrak{g}$ of exceptional type, there is a root $\alpha \in \Phi$ such that every irreducible submodule of $\bar{\rho}(\mathfrak{g})$ has a vector with nonzero $\mathfrak{l}_\alpha$ component and a vector with nonzero $\mathfrak{g}_{-\alpha}$ component.*

For our purpose, we only need to establish its analogs for $\mathfrak{g}$ of type A_n and B_n.

Lemma 3.19. *Assume l is large enough for $\mathfrak{g}$. For $\mathfrak{g}$ of type A_n, there is a root $\alpha \in \Phi$ such that every irreducible submodule of $\bar{\rho}(\mathfrak{g})$ has a vector with nonzero $\mathfrak{l}_\alpha$ component and a vector with nonzero $\mathfrak{g}_{-\alpha}$ component.*

Proof. Let $\mathfrak{g} = \mathfrak{sl}_{n+1}$ and let $\alpha_{i,j} = e_i - e_j$, $i \neq j$ be the roots of $\mathfrak{g}$. Let $E_{i,j}$ be the $n+1$ by $n+1$ matrix that has 1 at the (i, j)-entry and zeros elsewhere. The $\mathfrak{sl}_2$-triple associated to $\alpha_{i,j}$ is $\{X_{i,j} := E_{i,j}, H_{i,j} = E_{i,i} - E_{j,j}, Y_{i,j} := E_{j,i}\}$. Let

$$\begin{aligned} X &= X_{1,2} + X_{2,3} + \cdots + X_{n,n+1}, \\ H &= \sum_{i<j} H_{i,j} = k_1 H_{1,2} + k_2 H_{2,3} + \cdots + k_n H_{n,n+1}, \\ Y &= k_1 Y_{1,2} + k_2 Y_{2,3} + \cdots + k_n Y_{n,n+1}, \end{aligned}$$

where $k_i := i(n-i+1)$. The triple $\{X, H, Y\}$ is an $\mathfrak{sl}_2$-triple containing the regular unipotent element X.

A straightforward calculation gives for $i < j$

$$[Y, X_{i,j}] = k_i X_{i+1,j} - k_{j-1} X_{i,j-1}.$$

Put $h = j - i$ and apply the above identity recursively, we obtain

$$(adY)^h X_{i,j} = (-1)^h k_{i,j}\Big(H_{i,i+1} - \binom{h-1}{1} H_{i+1,i+2} + \binom{h-1}{2} H_{i+2,i+3} + \cdots + (-1)^{h-1}\binom{h-1}{h-1} H_{j-1}\Big),$$

where $k_{i,j} = k_i k_{i+1} \cdots k_{j-1}$. By Proposition 2.19,

$$\mathfrak{g}^X = \sum_{h=1}^{n} \langle v_{2h} \rangle$$

where $v_{2h} = \sum_{j-i=h} X_{i,j}$. Then we have

$$(adY)^h v_{2h} = h_1 H_{1,2} + \cdots + h_n H_{n,n+1}$$

with $h_1 = (-1)^h k_{1,h+1}$, $h_2 = (-1)^{h-1}\binom{h-1}{1} k_{1,h+1} + (-1)^h k_{2,h+2}$ and $h_i = (-1)^{h-1} h_{n-i+1}$. Since $(adY_{i,i+1})H_{i-1,i} = Y_{i,i+1}$, $(adY_{i,i+1})H_{i,i+1} = -2Y_{i,i+1}$ and $(adY_{i,i+1})H_{i+1,i+2} = Y_{i,i+1}$,

$$\begin{aligned} (adY)^{h+1} v_{2h} &= (adY)(h_1 H_{1,2} + \cdots + h_n H_{n,n+1}) \\ &= k_1(-2h_1 + h_2)Y_{1,2} + k_2(h_1 - 2h_2 + h_3)Y_{2,3} + \cdots + k_n(h_{n-1} - 2h_n)Y_{n,n+1}. \end{aligned}$$

One computes

$$h_2 - 2h_1 = (-1)^{h-1}(h+1)!h(n-1)(n-2)\cdots(n-h+1) \neq 0.$$

So if we let $\alpha = \alpha_{1,2}$ and suppose l is large enough for $\mathfrak{g}$, then the submodule of $\bar{\rho}(\mathfrak{g})$ generated by v_{2h} has a vector with nonzero $\mathfrak{g}_{-\alpha}$ component, that is, the vector $(adY)^{h+1} v_{2h}$; the vector $(adY)^h v_{2h}$ has nonzero $\mathfrak{l}_\alpha$ component, as $\alpha(h_1 H_{1,2} + \cdots + h_n H_{n,n+1}) = 2h_1 - h_2$ which is nonzero. $\square$

Corollary 3.20. *Assume l is large enough for $\mathfrak{g}$. For $\mathfrak{g}$ of type B_n, there is a root $\alpha \in \Phi$ such that every irreducible submodule of $\bar{\rho}(\mathfrak{g})$ has a vector with nonzero $\mathfrak{l}_\alpha$ component and a vector with nonzero $\mathfrak{g}_{-\alpha}$ component.*

Proof. Let $\mathfrak{g} = \mathfrak{so}_{2n+1}$. Let $V = k^{2n+1}$ be a vector space equipped with a bilinear form $x_1 y_{2n+1} + x_2 y_{2n} + \cdots + x_{2n+1} y_1$ with matrix J. Then $\mathfrak{g}$ can be identified with

$$\{X \in M_{2n+1}(k) \mid XJ + JX^t = 0\}.$$

The roots of $\mathfrak{g}$ are $e_i - e_j$, $e_i + e_j$, $-e_i - e_j$, $\pm e_i$ for $1 \le i \ne j \le n$. We choose a set of simple roots $\Delta = \{e_1 - e_2, e_2 - e_3, \cdots e_{n-1} - e_n, e_n\}$. The $\mathfrak{sl}_2$-triple associated to $e_i - e_{i+1}$ is $\{X_i := X_{i,i+1} - X_{2n-i+1,2n-i+2}, H_i := H_{i,i+1} + H_{2n-i+1,2n-i+2}, Y_i := Y_{i,i+1} - Y_{2n-i+1,2n-i+2}\}$, and the $\mathfrak{sl}_2$-triple associated to e_n is $\{X_n := X_{n,n+1} - X_{n+1,n+2}, H_n := 2H_{n,n+1} + 2H_{n+1,n+2}, Y_n := 2Y_{n,n+1} - 2Y_{n+1,n+2}\}$. Let

$$X = \sum_i X_i,$$
$$H = \sum_{1 \le i \le n-1} i(2n - i + 1) H_i + \tfrac{1}{2} n(n+1) H_n,$$
$$Y = \sum_{1 \le i \le n-1} i(2n - i + 1) Y_i + \tfrac{1}{2} n(n+1) Y_n.$$

A straightforward calculation shows that X, H, Y form an $\mathfrak{sl}_2$-triple containing the regular unipotent element $X \in \mathfrak{g}$.

Corresponding to the exponents $1, 3, \cdots, 2n-1$ of $\mathfrak{g}$, we put

$$v_{2\cdot 1} = X_{1,2} + \cdots + X_{n,n+1} - X_{n+1,n+2} - \cdots - X_{2n,2n+1} \in \mathfrak{so}_{2n+1},$$
$$v_{2\cdot 3} = X_{1,4} + \cdots + X_{n-1,n+2} - X_{n,n+3} - \cdots - X_{2n-2,2n+1} \in \mathfrak{so}_{2n+1}, \ldots, \ldots,$$
$$v_{2\cdot(2n-1)} = X_{1,2n} - X_{2,2n+1} \in \mathfrak{so}_{2n+1}.$$

Then $\mathfrak{g}^X = \sum_{i=1,3,\cdots,2n-1} < v_{2i} >$. Let $\alpha = e_1 - e_2$, the same calculation as in the proof of Lemma 3.19 gives $(adY)^i v_{2i}$ has a nonzero $\mathfrak{l}_\alpha$ component and $(adY)^{i+1} v_{2i}$ has a nonzero $\mathfrak{g}_{-\alpha}$ component for any exponent i. □

3D. ***Removing the congruence conditions on l.*** In this section, we use a result in [Fakhruddin et al. 2018] to remove the congruence condition we have imposed for G of type B_n, C_n, E_7 in Theorem 3.16. The following theorem is a simplified version of [Fakhruddin et al. 2018, Theorem 1.3]: as we are not considering geometric lifts, we relax the condition at l and only require the right hand side of Wiles' formula (Proposition 4.10) to be nonnegative. It applies to the residual representations we construct and allows us to deform it to an l-adic representation with Zariski-dense image for almost all primes l. However, their argument is very different and much more complicated than ours, so we only use it to remove the congruence conditions.

Theorem 3.21. *Suppose that there is a global deformation condition $\mathcal{L} = \{L_v\}_{v\in S}$ consisting of* smooth *local deformation conditions for each place $v \in S$. Let $K = \mathbb{Q}(\bar{\rho}(\mathfrak{g}), \mu_l)$. We assume the following*:

(1) $$\sum_{v\in S}(\dim L_v) \geq \sum_{v\in S} h^0(\Gamma_{\mathbb{Q}_v}, \bar{\rho}(\mathfrak{g})).$$

(2) *The field K does not contain μ_{l^2}.*

(3) *The groups $H^1(\mathrm{Gal}\,(K/\mathbb{Q}), \bar{\rho}(\mathfrak{g}))$ and $H^1(\mathrm{Gal}\,(K/\mathbb{Q}), \bar{\rho}(\mathfrak{g})(1))$ vanish.*

(4) *The spaces $\bar{\rho}(\mathfrak{g})$ and $\bar{\rho}(\mathfrak{g})(1)$ are semisimple $\mathbb{F}_l[\Gamma_{\mathbb{Q}}]$-modules (equivalently, $k[\Gamma_{\mathbb{Q}}]$-modules) having no common $\mathbb{F}_l[\Gamma_{\mathbb{Q}}]$-subquotient, and that neither contains the trivial representation.*

(5) *The space $\bar{\rho}(\mathfrak{g})$ is multiplicity-free as a $\mathbb{F}_l[\Gamma_{\mathbb{Q}}]$-module.*

Then there exists a finite set of primes Q disjoint from S, and a lift $\rho : \Gamma_{\mathbb{Q},S\cup Q} \to G(\mathcal{O})$ of $\bar{\rho}$ such that ρ is of type L_v at all $v \in S$.

Lemma 3.22. *Let $k = \mathbb{F}_l$ and let $\bar{\rho} : \Gamma_{\mathbb{Q}} \to G(k)$ be as in Sections 2A4–2A6. Then Theorem 3.21*(2)–(5) *hold.*

Proof. By the decomposition of $\bar{\rho}(\mathfrak{g})$ and the proof of Proposition 3.5(3)–(5) hold. It remains to show (2). Since by construction $\bar{\rho}$ is unramified at l, $\mathbb{Q}(\bar{\rho}(\mathfrak{g}))$ and $\mathbb{Q}(\mu_l)$ are linearly disjoint over $\mathbb{Q}$. It follows that

$$\mathrm{Gal}\,(K/\mathbb{Q}) \cong (\mathrm{Im}(\bar{\rho})/Z) \times (\mathbb{Z}/l\mathbb{Z})^\times$$

where Z denotes the center of $G(k)$. Assume K contains μ_{l^2}, then there would be a surjection

$$\mathrm{Gal}\,(K/\mathbb{Q})^{\mathrm{ab}} \twoheadrightarrow (\mathbb{Z}/l^2\mathbb{Z})^\times.$$

On the other hand, we have by the construction of $\bar{\rho}$ that $\mathrm{Im}(\bar{\rho})' = \mathrm{Im}(\bar{\rho})$ for G of type B_n and E_7, $\mathrm{Im}(\bar{\rho})'$ is of index two in $\mathrm{Im}(\bar{\rho})$ for G of type C_n. It follows that the order of $(\mathrm{Im}(\bar{\rho})/Z)^{\mathrm{ab}}$ is at most two, and hence the order of $\mathrm{Gal}\,(K/\mathbb{Q})^{\mathrm{ab}}$ is at most $2(l-1)$. But this is impossible since $(\mathbb{Z}/l^2\mathbb{Z})^\times$ has order $l(l-1)$ and $l \neq 2$. □

4. Simple, simply connected groups as monodromy groups

In this section, we prove Theorem 1.2 for G a simple, simply connected algebraic group. Recall that there are two different constructions for the residual representation $\bar{\rho} : \Gamma_{\mathbb{Q}} \to G(k)$: one has image a large index subgroup of $N_G(T)_k$ with the properties that $\bar{\rho}$ is unramified at l and $\mathrm{ad}\,\bar{\rho}(c)$ is nontrivial; the other factors through a principal GL_2 such that $\bar{\rho}(c) = \rho^\vee(-1)$.

4A. *Local deformation conditions.* We need to define several local deformation conditions for deforming the mod p representations.

4A1. *The archimedean place.* Recall that in Sections 2A3–2A5, we construct the residual representations by first realizing S_n or A_n as a Galois group over $\mathbb{Q}$ and then repeatedly applying Theorem 2.5 to build the Galois extension realizing N or a subgroup of it over $\mathbb{Q}$. We write c for the nontrivial element in $\Gamma_\mathbb{R}$, the complex conjugation.

Proposition 4.1. *Let G be of classical type and $\bar{\rho} : \Gamma_\mathbb{Q} \to G(k)$ be as in Sections 2A2-2A5. In particular,* ad $\bar{\rho}(c)$ *is nontrivial. Then*:

(1) *For G of type A_{n-1}, $h^0(\Gamma_\mathbb{R}, \bar{\rho}(\mathfrak{g})) \le n^2 - 2n + 1$.*

(2) *For G of type B_n, $h^0(\Gamma_\mathbb{R}, \bar{\rho}(\mathfrak{g})) \le 2n^2 - 3n + 2$.*

(3) *For G of type C_n, $h^0(\Gamma_\mathbb{R}, \bar{\rho}(\mathfrak{g})) \le 2n^2 - 3n + 4$.*

(4) *For G of type D_n, $h^0(\Gamma_\mathbb{R}, \bar{\rho}(\mathfrak{g})) \le 2n^2 - 5n + 4$.*

Proof. Let f be the number of root vectors that are fixed by $\bar{\rho}(c)$. Let us recall that ad $\bar{\rho}(c)$ is nontrivial by Theorem 2.4.

If G is of type A_{n-1}, we have $G(k) \cong \mathrm{SL}(V)$ with $V = k^n$. Let $d := \dim V^{\bar{\rho}(c)}$. Then

$$f = 2\left(\binom{d}{2} + \binom{n-d}{2}\right).$$

If G is of type B_n, $G(k)/\mu_2 \cong \mathrm{SO}(V)$ with $V = k^{2n+1}$ a k-vector space equipped with the nondegenerate symmetric bilinear form $x_1 y_{2n+1} + x_2 y_{2n} + \cdots + x_{2n+1} y_1$. We may assume $\bar{\rho}(c)$ is conjugate to $\mathrm{diag}(\epsilon_1, \ldots, \epsilon_n, 1, \epsilon_n, \ldots, \epsilon_1)$ in $\mathrm{SO}(V)$ and let d be the number of ones among $\epsilon_1, \ldots, \epsilon_n$. Then

$$f = 4\left(\binom{d}{2} + \binom{n-d}{2}\right) + 2d.$$

If G is of type C_n, $G(k) \cong \mathrm{Sp}(V)$ with $V = k^{2n}$ a k-vector space equipped with the nondegenerate alternating bilinear form $x_1 y_{2n} + \cdots x_n y_{n+1} - x_{n+1} y_n - \cdots - x_{2n} y_1$. We may assume $\bar{\rho}(c)$ is conjugate to $\mathrm{diag}(\epsilon_1, \ldots, \epsilon_n, \epsilon_n, \ldots, \epsilon_1)$ and let d be the number of 1's among $\epsilon_1, \ldots, \epsilon_n$. Then

$$f = 4\left(\binom{d}{2} + \binom{n-d}{2}\right) + 2n.$$

If G is of type D_n, a quotient of $G(k)$ is isomorphic to $\mathrm{SO}(V)$ with $V = k^{2n}$ a k-vector space equipped with the nondegenerate symmetric bilinear form $x_1 y_{2n} + x_2 y_{2n-1} + \cdots + x_{2n} y_1$. We may assume $\bar{\rho}(c)$ is conjugate to $\mathrm{diag}(\epsilon_1, \ldots, \epsilon_n, \epsilon_n, \ldots, \epsilon_1)$ in $\mathrm{SO}(V)$ and let d be the number of 1's among $\epsilon_1, \ldots, \epsilon_n$. Then

$$f = 4\left(\binom{d}{2} + \binom{n-d}{2}\right).$$

By the construction of $\bar{\rho}$ (see the paragraph before Remark 2.8 in Section 2A3 for type A_n; for other types, see Sections 2A4, 2A5 and 2A6), ad $\bar{\rho}(c)$ is nontrivial, which implies $0 < d < n$. So f attains its maximum when $d = n - 1$. Since

$$h^0(\Gamma_\mathbb{R}, \bar{\rho}(\mathfrak{g})) = rk(\mathfrak{g}) + f,$$

the upper bounds can then be computed easily. □

Proposition 4.2. *Let G be of type E_7 and $\bar\rho : \Gamma_{\mathbb{Q}} \to G(k)$ be as in Section 2A6. In particular*, ad $\bar\rho(c)$ *is nontrivial. Then* $h^0(\Gamma_{\mathbb{R}}, \bar\rho(\mathfrak{g})) \le 7 + 126 - 14 = 119$.

Proof. Suppose that $\bar\rho(c) \in T(k)$ for a maximal torus T split over k. Let $\Phi = \Phi(G, T)$, and $\mathfrak{g}_\Phi = \sum_{\alpha\in\Phi} \mathfrak{g}_\alpha$ be the k-subspace of $\bar\rho(\mathfrak{g})$ generated by all root vectors. The Lie algebra $\mathfrak{t}$ of $T(k)$, which has k-dimension 7, is clearly fixed by ad $\bar\rho(c)$. Thus, it suffices to show that the -1-eigenspace of ad $\bar\rho(c) \mid \mathfrak{g}_\Phi$ has k-dimension at least 14. We consider ad $\bar\rho(c) \mid \mathfrak{g}_{\Phi'}$ where $\Phi' \subset \Phi$ is of type A_7. This action is nontrivial. By the A_n-calculation in the proof of Proposition 4.1 (letting $n = 7$), the -1-eigenspace of ad $\bar\rho(c) \mid \mathfrak{g}_{\Phi'}$ has dimension at least twice the rank of Φ', proves the proposition. □

The following lemma is clear.

Lemma 4.3. *The dimension*

$$\dim_k(\mathrm{Sym}^{2n}(k^2) \otimes \det^{-n})^{\mathrm{diag}(1,-1)}$$

equals n when n is odd, and $n+1$ when n is even.

Corollary 4.4. *Let $\bar\rho : \Gamma_{\mathbb{Q}} \to G(k)$ be as in Section 2B. Then*

$$h^0(\Gamma_{\mathbb{R}}, \bar\rho(\mathfrak{g})) = 4 \qquad \text{for } G = \mathrm{SL}_3,$$
$$h^0(\Gamma_{\mathbb{R}}, \bar\rho(\mathfrak{g})) = 9 \qquad \text{for } G = \mathrm{Spin}_7,$$
$$h^0(\Gamma_{\mathbb{R}}, \bar\rho(\mathfrak{g})) = 38 \qquad \text{for } G = E_6^{\mathrm{sc}}.$$

Proof. This follows from Lemma 4.3 and Proposition 2.19. Note that the complex conjugation maps to diag$(1, -1)$ in GL_2 because the representation $r_{f,\lambda}$ in the last paragraph of Section 2 is odd. □

4A2. *The place l.* As we are not looking for geometric l-adic Galois representations in this paper, we impose *no condition* at the place l. So the tangent space is $H^1(\Gamma_l, \bar\rho(\mathfrak{g}))$. By the local Euler characteristic formula,

$$h^1(\Gamma_l, \bar\rho(\mathfrak{g})) = h^0(\Gamma_l, \bar\rho(\mathfrak{g})) + h^2(\Gamma_l, \bar\rho(\mathfrak{g})) + \dim_k \mathfrak{g}.$$

Lemma 4.5. *Let $\bar\rho$ be as in Section 2A or 2B. Then $h^2(\Gamma_l, \bar\rho(\mathfrak{g})) = 0$ for large enough primes l.*

Proof. By local duality, it suffices to show that $h^0(\Gamma_l, \bar\rho(\mathfrak{g})(1)) = 0$. For the representation $\bar\rho$ in Section 2A, $\bar\rho(I_{\mathbb{Q}_l})$ is trivial by construction but $\bar\kappa(I_{\mathbb{Q}_l})$ is nontrivial, so $\bar\rho(\mathfrak{g})(1)^{I_{\mathbb{Q}_v}}$ is trivial. In particular, $h^0(\Gamma_l, \bar\rho(\mathfrak{g})(1)) = 0$. For the representation $\bar\rho$ in Section 2B, Proposition 2.19 and the lemma below imply $h^0(\Gamma_l, \bar\rho(\mathfrak{g})(1)) = 0$. □

Lemma 4.6. *We have $h^0(\Gamma_l, \mathrm{Sym}^{2m}(\bar r_f) \otimes \det(\bar r_f)^{-m} \otimes \bar\kappa) = 0$ for $m \ge 1$ and large enough primes l (relative to m).*

Proof. The argument is similar to the proof of [Weston 2004, Proposition 4.4]. Let K be a finite extension of $\mathbb{Q}_l$ with ring of integers $\mathcal{O}$ and residue field k. Let $v : K^\times \to \mathbb{Q}$ be the valuation on K, normalized so that $v(l) = 1$. We briefly recall the setting in Section 4.1 of [Weston 2004]. For $a < b$, let $\mathcal{MF}^{a,b}(\mathcal{O})$ denote the category of filtered Dieudonné $\mathcal{O}$-modules D equipped with a decreasing filtration of $\mathcal{O}$-modules $\{D_i\}_{i\in\mathbb{Z}}$ and a family of $\mathcal{O}$-linear maps $\{f_i : D^i \to D\}$ satisfying $D^a = D$ and $D^b = 0$; see [Weston 2004, Definition 4.1]. Let $\mathcal{G}^{a,b}(\mathcal{O})$ denote the category of finite type $\mathcal{O}$-module subquotients of crystalline K-representations V with $D^a_{\mathrm{crys}}(V) = D_{\mathrm{crys}}(V)$ and $D^b_{\mathrm{crys}}(V) = 0$. Fontaine–Laffaille define a functor

$$\mathcal{U} : \mathcal{MF}^{a,a+l}(\mathcal{O}) \to \mathcal{G}^{a,a+l}(\mathcal{O})$$

satisfying a list of properties including $\mathcal{U}$ that is stable under formation of subobjects and quotients, and compatible with tensor products for l large enough. In particular, $\mathcal{U}$ is compatible with symmetric powers for large enough l. Let $\varepsilon : \Gamma_l \to \mathcal{O}^\times$ be an unramified character of finite order and let $\mathcal{O}(\varepsilon)$ denote a free $\mathcal{O}$-module of rank 1 with Γ_l-action via ε. Then $\mathcal{O}(\varepsilon) \in \mathcal{G}^{0,1}(\mathcal{O})$, so that there is $D_\varepsilon \in \mathcal{MF}^{0,1}(\mathcal{O})$ such that $\mathcal{U}(D_\varepsilon) = \mathcal{O}(\varepsilon)$. This D_ε is a free $\mathcal{O}$-module of rank one with $D_\varepsilon = D^0_\varepsilon$ and $f_0 : D^0_\varepsilon \to D_\varepsilon$ the multiplication by $\varepsilon^{-1}(l) := \varepsilon^{-1}(\mathrm{Fr}_l)$ where the Frobenius element is arithmetic.

Let f be a newform of weight 3, level N, and character ε. Let $K \supset E_\lambda$ and let $r_f := r_{f,\lambda} : \Gamma_{\mathbb{Q}} \to \mathrm{GL}_2(K)$ be the Galois representation associated to f of weight k (see Example 1.5). We fix an embedding $\Gamma_l \to \Gamma_{\mathbb{Q}}$, and let V_f be a two dimensional K-vector space on which Γ_l acts via $r_f \mid \Gamma_l$, and fix a Γ_l-stable $\mathcal{O}$-lattice $T_f \subset V_f$. If l does not divide N, then V_f is crystalline and $T_f \in \mathcal{G}^{0,3}(\mathcal{O})$. Thus for $l > k$ there exists $D_f \in \mathcal{MF}^{0,3}(\mathcal{O})$ with $\mathcal{U}(D_f) \cong T_f$. The filtration on D_f satisfies $\mathrm{rk}_{\mathcal{O}}(D^i_f) = 2$ if $i \le 0$, $\mathrm{rk}_{\mathcal{O}}(D^i_f) = 1$ if $1 \le i \le 2$, and $\mathrm{rk}_{\mathcal{O}}(D^i_f) = 0$ if $i \ge 3$. Choose an $\mathcal{O}$-basis x, y of D_f with x an $\mathcal{O}$-generator of D^1_f. Let $a, b, c, d \in \mathcal{O}$ be such that $f_0 x = ax + by$, $f_0 y = cx + dy$. Then $a + d = a_l$ and $ad - bc = l^2\varepsilon(l)$. We have $v(a), v(b) \ge 2$.

Let $\bar{r}_f : \Gamma_l \to \mathrm{GL}_2(k)$ be the Galois representation $T_f/\lambda T_f$. Since $\det r_f = \kappa^2\varepsilon$, we have

$$\mathrm{Sym}^{2m}(T_f) \otimes \det(T_f)^{-m} \otimes \kappa = (\mathrm{Sym}^{2m}(T_f) \otimes \mathcal{O}(\varepsilon^{-m}))(1 - 2m).$$

When l is large enough relative to m, by [Fontaine and Messing 1987, Proposition 1.7] we can take

$$D = (\mathrm{Sym}^{2m}(D_f) \otimes D_{\varepsilon^{-m}})(2m - 1).$$

Further, since $(\mathrm{Sym}^{2m}(T_f) \otimes \det(T_f)^{-m} \otimes \kappa)/\lambda$ is a realization of $\mathrm{Sym}^{2m}(\bar{r}_f) \otimes \det(\bar{r}_f)^{-m} \otimes \bar{\kappa}$, we have

$$H^0(\Gamma_l, \mathrm{Sym}^{2m}(\bar{r}_f) \otimes \det(\bar{r}_f)^{-m} \otimes \bar{\kappa}) = \ker(1 - f_0 : D^0/\lambda D^0 \to D/\lambda D).$$

By the definition of Tate twists and tensor products of filtered Dieudonné $\mathcal{O}$-modules, we have

$$D^0 = (\mathrm{Sym}^{2m}(D_f) \otimes D_{\varepsilon^{-m}})^{2m-1} = \sum_{i_1+\cdots+i_{2m}+j=2m-1} D^{i_1}_f \cdots D^{i_{2m}}_f \cdot D^j_{\varepsilon^{-m}} = \sum_{i_1+\cdots+i_{2m}=2m-1} D^{i_1}_f \cdots D^{i_{2m}}_f \cdot D^0_{\varepsilon^{-m}}.$$

To make the sum nonzero, there must be at least m indices that are greater than or equal to 1, so at least m indices must be two since $x \in D^2_f$ as well as D^1_f. It follows that $\{x^i y^{2m-i-1} w \mid i \ge m\}$ is an $\mathcal{O}$-basis of D^0, where w is an $\mathcal{O}$-generator of $D_{\varepsilon^{-m}}$. We compute

$$f_0(x^i y^{2m-i-1} w) = \frac{\varepsilon^m(l)}{l^{2m-1}}(ax + by)^i (cx + dy)^{2m-i-1}.$$

Since $v(a), v(b) \geq 2$ and $i \geq m$, all the coefficients of x and y have positive valuations. Therefore, $f_0(x^i y^{2m-i-1} w) \equiv 0$ modulo λ, which implies $f_0 : D^0/\lambda D^0 \to D/\lambda D$ is zero. It follows that $H^0(\Gamma_l, \mathrm{Sym}^{2m}(\bar{r}_f) \otimes \det(\bar{r}_f)^{-m} \otimes \bar{\kappa})$ is trivial. □

Corollary 4.7. $h^1(\Gamma_l, \bar{\rho}(\mathfrak{g})) = h^0(\Gamma_l, \bar{\rho}(\mathfrak{g})) + \dim_k \mathfrak{g}$.

4A3. *A zero-dimensional deformation.* In order to maximize the Zariski-closure of the image of the l-adic lift of the residual representation, we need to impose a simple local deformation condition at some unramified place.

Suppose that $p \neq l$, F is a finite extension of $\mathbb{Q}_p$, and $\bar{\rho} : \Gamma_F \to G(k)$ is an unramified representation. Let $g \in G(\mathcal{O})$ be a lift of $\bar{\rho}(\mathrm{Fr}_p)$.

Definition 4.8. Define

$$\mathrm{Lift}_{\bar{\rho}}^{g} : \mathrm{CNL}_{\mathcal{O}} \to \textbf{Sets}$$

such that for a complete local noetherian $\mathcal{O}$-algebra R, $\mathrm{Lift}_{\bar{\rho}}^{g}(R)$ consists of all lifts

$$\rho : \Gamma_F \to G(R)$$

of $\bar{\rho}$ such that ρ is unramified and $\rho(\mathrm{Fr}_p)$ is $\hat{G}(R)$-conjugate to g.

So the tangent space is zero-dimensional and when $\mathrm{Lift}_{\bar{\rho}}^{g}$ is a local deformation condition, it is clearly smooth. But for a given g, $\mathrm{Lift}_{\bar{\rho}}^{g}$ may not be representable. But at least we have

Proposition 4.9. *Suppose that G is simply connected. Let $\bar{g}$ and g be regular semisimple elements of $G(\bar{\mathbb{F}}_l)$ and $G(\bar{\mathbb{Q}}_l)$, respectively. Then $\mathrm{Lift}_{\bar{\rho}}^{g}$ is representable.*

Proof. By Schlessinger's criterion, it suffices to show the following: for any $A \twoheadrightarrow B$ in $\mathrm{CLN}_{\mathcal{O}}$ with kernel I for which $I \cdot \mathfrak{m}_A = 0$, the induced map

$$Z_G(g)(A) \to Z_G(g)(B)$$

is surjective. The group $Z_G(g)$ is a scheme over $\mathcal{O}$, we denote the structure map by

$$f : Z_G(g) \to \mathrm{Spec}\, \mathcal{O}.$$

We need to show that $Z_G(g)$ is a smooth $\mathcal{O}$-scheme. It suffices to show that:

- The map f is flat over $\mathcal{O}$.
- The generic fiber and the special fiber of f are smooth of the same dimension.

Because G is simply connected and $\bar{g}$ and g are regular semisimple, $Z_G(g)(\mathcal{O}/\lambda)$ and $Z_G(g)(\mathrm{Frac}\,\mathcal{O})$ are connected maximal tori of $G(\mathcal{O}/\lambda)$ and $G(\mathrm{Frac}\,\mathcal{O})$, respectively, with dimension the rank of G. The second bullet follows.

To show the first bullet, note that f has a section, that is, $Z_G(g)$ has an $\mathcal{O}$-point (for example, the element $g \in G(\mathcal{O})$ itself). Moreover, by the previous paragraph, the generic fiber and the special fiber of f are both irreducible, reduced and have the same dimension. It follows from Proposition 6.1 of [Gan and Yu 2003] that f is flat. □

4A4. *Steinberg deformations.* In [Patrikis 2016, Section 4.3], a local deformation condition of "Steinberg type" is taken at a place in order to obtain a regular unipotent element in the image of the l-adic lift. *We will only need this in deforming those $\bar{\rho}$ constructed from the principal* GL_2. We refer the reader to [Patrikis 2016, Section 4.3] for the definition and properties of the Steinberg deformation condition. The dimension of the tangent space equals $h^0(\Gamma_v, \bar{\rho}(\mathfrak{g}))$.

4A5. *Minimal prime to l deformations.* This deformation condition is well known; see [Patrikis 2016, Section 4.4] for its definition. We will use this deformation condition at places $v \neq l$ for which $\bar{\rho}(I_{\mathbb{Q}_v})$ is nontrivial and $\bar{\rho}(\Gamma_{\mathbb{Q}_v})$ has order prime to l. The tangent space is $H^1(\Gamma_v/I_v, \bar{\rho}(\mathfrak{g})^{I_v})$, whose dimension is $h^0(\Gamma_v, \bar{\rho}(\mathfrak{g}))$.

4B. ***Deforming* mod *p Galois representations.*** In this section, we specify the global deformation condition and compute the Wiles formula, then use the results in Section 3 to prove Theorem 1.2. Let us recall Wiles' formula, for a proof; see [Patrikis 2016, Proposition 9.2].

Proposition 4.10. *Let M be a finite-dimensional k-vector space with a continuous $\Gamma_{\mathbb{Q}}$ action unramified outside a finite set of places S. Let $\mathcal{L} = \{L_v\}_{v\in S}$ and $\mathcal{L}^\perp = \{L_v^\perp\}_{v\in S}$ be a Selmer system and dual Selmer system, respectively, for M. Then*

$$h^1_{\mathcal{L}}(\Gamma_S, M) - h^1_{\mathcal{L}^\perp}(\Gamma_S, M^\vee) = h^0(\Gamma_S, M) - h^0(\Gamma_S, M^\vee) + \sum_{v\in S}(\dim_k L_v - h^0(\Gamma_v, M)).$$

We will compute the right-hand side of the identity for $M = \bar{\rho}(\mathfrak{g})$ or $\bar{\rho}(\mathfrak{t})$ and for a global deformation condition to be specified below. For $\bar{\rho}(\mathfrak{g})$ from either Section 2A or 2B, note that $h^0(\Gamma_S, \bar{\rho}(\mathfrak{g})) = h^0(\Gamma_S, \bar{\rho}(\mathfrak{g})(1)) = 0$.

4B1. *Weyl group case.* For G a simple, simply connected group of classical type or type E_7, let $\bar{\rho} : \Gamma_{\mathbb{Q}} \to G(k)$ be as in Section 2A. Here we exclude the A_1, A_2, B_3 cases. We impose no condition at $v = l$ which is liftable by Lemma 4.5, and impose the minimal prime to l condition at $v \in S - \{\infty, l\}$ (note that $\bar{\rho}(\Gamma_{\mathbb{Q}})$ has order prime to l by our construction). Moreover, we will find a prime $p \notin S$ for which $\bar{\rho}(\mathrm{Fr}_p)$ is regular semisimple, together with a regular semisimple lift $g \in G(\mathcal{O})$ of $\bar{\rho}(\mathrm{Fr}_p)$. Then we take the deformation condition $\mathrm{Lift}^g_{\bar{\rho}|_{\Gamma_p}}$ at p.

Lemma 4.11.
$$\sum_{v\in S} \dim_k L_v \geq \sum_{v\in S} h^0(\Gamma_v, \bar{\rho}(\mathfrak{g})).$$

Proof. This follows directly the local computations in Section 4A, for example Proposition 4.1 and Corollary 4.7, etc. We record here a lower bound for $\sum_{v\in S} \dim_k L_v - \sum_{v\in S} h^0(\Gamma_v, \bar{\rho}(\mathfrak{g}))$. For $G = \mathrm{SL}_n$, it is $n-1$; for $G = \mathrm{Spin}_{2n+1}$, it is $3n-2$; for $G = \mathrm{Sp}_{2n}$, it is $3n-4$; for $G = \mathrm{Spin}_{2n}$, it is $3n-4$; and for $G = E_7$, it is 7. □

Lemma 4.12.
$$\sum_{v\in S} \dim_k(L_v \cap H^1(\Gamma_v, \bar{\rho}(\mathfrak{t}))) \leq \sum_{v\in S} h^0(\Gamma_v, \bar{\rho}(\mathfrak{t})).$$

Proof. For $v \notin \{\infty, l, p\}$, L_v corresponds to the minimal prime to l deformation condition and we have

$$\dim_k(L_v \cap H^1(\Gamma_v, \bar{\rho}(\mathfrak{t}))) = \dim_k H^1(\Gamma_v/I_v, \bar{\rho}(\mathfrak{t})^{I_v}) = h^0(\Gamma_v, \bar{\rho}(\mathfrak{t})).$$

So it suffices to compare both sides for $v \in \{\infty, l, p\}$. The left-hand side subtracting the right-hand side equals

$$(0 - h^0(\Gamma_{\mathbb{R}}, \bar{\rho}(\mathfrak{t}))) + (h^1(\Gamma_l, \bar{\rho}(\mathfrak{t})) - h^0(\Gamma_l, \bar{\rho}(\mathfrak{t}))) + (0 - h^0(\Gamma_p, \bar{\rho}(\mathfrak{t}))).$$

By local duality and Lemma 4.5,

$$h^1(\Gamma_l, \bar{\rho}(\mathfrak{t})) - h^0(\Gamma_l, \bar{\rho}(\mathfrak{t})) = \dim_k \mathfrak{t}.$$

Combining this with the identity

$$h^0(\Gamma_p, \bar{\rho}(\mathfrak{t})) = \dim_k \mathfrak{t},$$

we see that the difference is $-h^0(\Gamma_{\mathbb{R}}, \bar{\rho}(\mathfrak{t})) \leq 0$. □

Let us make the following observation which is from [Patrikis 2016, Lemma 7.7]. It will be used frequently in the proof of Propositions 4.14, 4.18, and 4.20; suppose that $\bar{\rho} : \Gamma_{\mathbb{Q}} \to G(k)$ is a continuous representation with a continuous lift $\rho : \Gamma_{\mathbb{Q}} \to G(\mathcal{O})$. Let G_ρ be the Zariski closure of $G(\mathcal{O})$ in $G(\bar{\mathbb{Q}}_l)$. Then $\mathrm{Lie}(G_\rho)$, $\mathrm{Lie}(G_\rho) \cap \mathfrak{g}_{\mathcal{O}}$, and $(\mathrm{Lie}(G_\rho) \cap \mathfrak{g}_{\mathcal{O}}) \otimes_{\mathcal{O}} k$ are $\Gamma_{\mathbb{Q}}$-modules. Moreover, the last one is a submodule of $\bar{\rho}(\mathfrak{g})$ and thus is a direct sum of some irreducible summands of $\bar{\rho}(\mathfrak{g})$. If $\mathrm{Lie}(G_\rho) = \mathfrak{g}(\bar{\mathbb{Q}}_l)$ (which is equivalent to $(\mathrm{Lie}(G_\rho) \cap \mathfrak{g}_{\mathcal{O}}) \otimes_{\mathcal{O}} k = \mathfrak{g}$), then $G_\rho = G(\bar{\mathbb{Q}}_l)$ (since G is connected).

Lemma 4.13. *Let G be a semisimple algebraic group defined over $\mathcal{O}$ and let C be a proper subvariety of G. Let $g \in G(\mathcal{O})$ and $H = g\hat{G}(\mathcal{O})$ be the corresponding $\hat{G}(\mathcal{O})$-coset of $G(\mathcal{O})$. Then there is a regular semisimple element in $H - C(\bar{\mathbb{Q}}_l)$.*

Proof. Let V be the union of $C(\bar{\mathbb{Q}}_l)$ and the set of elements of $G(\bar{\mathbb{Q}}_l)$ that are not regular semisimple. Then V is a proper Zariski-closed subset of $G(\bar{\mathbb{Q}}_l)$ as the set of regular semisimple elements is Zariski-open (see for example [Humphreys 1995, Theorem 2.5]). On the other hand, H is Zariski-dense in $G(\bar{\mathbb{Q}}_l)$. If there were no regular semisimple element in $H - C(\bar{\mathbb{Q}}_l)$, then $H \subset V$ and H could not be Zariski-dense, a contradiction. □

Proposition 4.14. *Let G be a simple, simply connected group of classical type (excluding type A_1, A_2 and B_3) or type E_7. Then for almost all primes l when Φ is of type A_n or D_n, for almost all primes $l \equiv 1(4)$ when Φ is of type B_n or C_n, and for almost all primes $l \equiv 1(3)$ when Φ is of type E_7, there are l-adic lifts*

$$\rho : \Gamma_{\mathbb{Q}} \to G(\mathcal{O})$$

of $\bar{\rho} : \Gamma_{\mathbb{Q}} \to G(k)$ defined in Section 2A with Zariski-dense image in $G(\bar{\mathbb{Q}}_l)$.

Proof. By Lemmas 4.11 and 4.12, we can apply Theorem 3.16 to obtain a lift $\rho : \Gamma_{\mathbb{Q}} \to G(\mathcal{O})$ satisfying the prescribed local conditions. The condition at p implies that G_ρ has infinitely many elements and so $\mathrm{Lie}(G_\rho)$ is nontrivial.

If G is of type A_n or D_n, by Propositions 2.9 and 2.13, $(\mathrm{Lie}(G_\rho)\cap\mathfrak{g}_\mathcal{O})\otimes_\mathcal{O} k$ (as a Lie subalgebra of $\mathfrak{g}$) is then either $\mathfrak{t}$ or $\mathfrak{g}$. By the previous lemma, there exist a regular semisimple element $g\in G(\mathcal{O})$ such that $\bar{g}=\bar{\rho}(\mathrm{Fr}_p)$ and $g\notin N_G(T)(\bar{\mathbb{Q}}_l)$. Imposing $\mathrm{Lift}^g_{\bar{\rho}\,|\,p}$ at p, we obtain $G_\rho\not\subseteq N_G(T)(\bar{\mathbb{Q}}_l)$, which implies $(\mathrm{Lie}(G_\rho)\cap\mathfrak{g}_\mathcal{O})\otimes_\mathcal{O} k$ cannot be $\mathfrak{t}$.

If G is of type B_n, by Proposition 2.13, $(\mathrm{Lie}(G_\rho)\cap\mathfrak{g}_\mathcal{O})\otimes_\mathcal{O} k$ (as a Lie subalgebra of $\mathfrak{g}$) is $\mathfrak{t}$, $\mathfrak{t}\oplus\mathfrak{g}_l\cong\mathfrak{so}_{2n}$ or $\mathfrak{g}$. Let H be an algebraic subgroup of G containing T such that $(\mathrm{Lie}(H)\cap\mathfrak{g}_\mathcal{O})\otimes_\mathcal{O} k=\mathfrak{t}\oplus\mathfrak{g}_l$ (there are finitely many of them). Let C be the union of $N_G(T)$ and all such H, which is a proper subvariety of G. By the previous lemma, there exist a regular semisimple element $g\in G(\mathcal{O})$ such that $\bar{g}=\bar{\rho}(\mathrm{Fr}_p)$ and $g\notin C(\bar{\mathbb{Q}}_l)$. Imposing $\mathrm{Lift}^g_{\bar{\rho}\,|\,p}$ at p, we obtain $G_\rho\not\subseteq C(\bar{\mathbb{Q}}_l)$, which implies $(\mathrm{Lie}(G_\rho)\cap\mathfrak{g}_\mathcal{O})\otimes_\mathcal{O} k$ cannot be $\mathfrak{t}$ or $\mathfrak{t}\oplus\mathfrak{g}_l$.

If G is of type C_n, by Proposition 2.11, $(Lie(G_\rho)\cap\mathfrak{g}_\mathcal{O})\otimes_\mathcal{O} k$ (as a Lie subalgebra of $\mathfrak{g}$) is $\mathfrak{t}$, $\mathfrak{t}\oplus\mathfrak{g}_l\cong(\mathfrak{sl}_2)^n$ or $\mathfrak{g}$. The same argument as for type B_n enables us to impose a suitable condition $\mathrm{Lift}^g_{\bar{\rho}\,|\,p}$ at p in order to force $(\mathrm{Lie}(G_\rho)\cap\mathfrak{g}_\mathcal{O})\otimes_\mathcal{O} k=\mathfrak{g}$.

Finally, suppose that $G=E_7^{\mathrm{sc}}$. By Proposition 2.17, $(Lie(G_\rho)\cap\mathfrak{g}_\mathcal{O})\otimes_\mathcal{O} k$ (as a Lie subalgebra of $\mathfrak{g}$) is either $\mathfrak{t}$, $\mathfrak{t}\oplus\mathfrak{g}_a\cong\mathfrak{sl}_7$ or $\mathfrak{g}$. We can force $(\mathrm{Lie}(G_\rho)\cap\mathfrak{g}_\mathcal{O})\otimes_\mathcal{O} k$ to be $\mathfrak{g}$ in the same way as above. □

Remark 4.15. As we have flexibilities in choosing $g\in G(\mathcal{O})$ lifting $\bar{\rho}(\mathrm{Fr}_p)$, it is easy to see that there are infinitely many lifts ρ that are nonconjugate in G^{ad}.

4B2. *Principal* GL_2 *case.* For G a simply connected group of one of the following types: A_2, B_3, E_6, let $\bar{\rho}:\Gamma_\mathbb{Q}\to G(k)$ be as in Section 2B.

We begin with the following proposition due to Tom Weston [2004, Proposition 5.3].

Proposition 4.16. *Let $\pi=\pi_f$ be a cuspidal automorphic representation corresponding to a holomorphic eigenform f of weight at least 2. Assume that for some prime p, π_p is isomorphic to a twist of the Steinberg representation of $\mathrm{GL}_2(\mathbb{Q}_p)$. Then for almost all λ, the local Galois representation $\bar{r}_{f,\lambda}|_{\Gamma_p}$ (in the notation of Example 1.5) has the form*

$$\bar{r}_{f,\lambda}|_{\Gamma_p}\sim\begin{pmatrix}\chi\bar{\kappa} & *\\ 0 & \chi\end{pmatrix}$$

where the extension $*$ in $H^1(\Gamma_p,k_\lambda(\bar{\kappa}))$ is nonzero.

Let f be a non-CM weight 3 cuspidal eigenform that is a newform of level $\Gamma_1(p)\cap\Gamma_0(q)$ for some primes p and q; the nebentypus of f is a character $\varepsilon:(\mathbb{Z}/pq\mathbb{Z})^\times\to(\mathbb{Z}/p\mathbb{Z})^\times\to\mathbb{C}^\times$. Such a form f exists, see for example, [LMFDB 2013, 15.3.7.a and 15.3.11.a]. Note that the automorphic representation π_f associated to f is Steinberg at p. Proposition 4.16 together with the definition of principal GL_2 then imply $\bar{\rho}|_{\Gamma_p}$ is Steinberg in the sense of [Patrikis 2016, Definition 4.13]. At p we take the Steinberg deformation condition. As π_f is a principal series at q, $\bar{\rho}(I_q)$ has order prime to l. We then use the minimal prime to l deformation condition at q. At l we impose no condition. Moreover, choose an element $\sigma\in\Gamma_\mathbb{Q}$ such that $\bar{\rho}(\sigma)$ is regular semisimple in $T(k)$ together with a lift $g\in T(\mathcal{O})$ such that $\alpha(g)$,

$\alpha \in \Delta$ are distinct. By Chebotarev's density theorem, there is a prime $r \notin \{\infty, l, p, q\}$ such that $\mathrm{Fr}_r = \sigma$. We then take $\mathrm{Lift}^{g}_{\bar{\rho}|_{\Gamma_r}}$ at r. Let $\mathcal{L}$ be the Selmer system associated to the above local deformations.

Lemma 4.17. *The right-hand side of Wiles' formula is* 2 *for* A_2, 9 *for* B_3, *and* 34 *for* E_6.

Proof. This follows from Corollaries 4.4, 4.7 and Proposition 4.10. □

Proposition 4.18. *For* $G = \mathrm{SL}_3$, Spin_7 *or* E_6^{sc} *and for almost all primes* l, *there are* l*-adic lifts*

$$\rho : \Gamma_{\mathbb{Q}} \to G(\mathcal{O})$$

of $\bar{\rho} : \Gamma_{\mathbb{Q}} \to G(k)$ *defined as in Section 2B with Zariski-dense image in* $G(\bar{\mathbb{Q}}_l)$.

Proof. The proof is very similar to the proof of [Patrikis 2016, Theorem 7.4], so we skip a few details here.

We first show that Theorem 3.4 applies to $\bar{\rho}$. Item (1) is satisfied by Lemma 4.17; items (2) and (3) are satisfied by the proof of [Patrikis 2016, Theorem 7.4]; for item (4), take $\sigma \in \Gamma_{\mathbb{Q}}$ such that $\bar{\rho}(\sigma) = 2\rho^{\vee}(a)$ is regular with $1 \neq a \in (\mathbb{Z}/l\mathbb{Z})^{\times}$ and $\bar{\kappa}(\sigma) = a^2$ (which is possible, again, see the proof of [Patrikis 2016, Theorem 7.4]. It follows that item (a) is satisfied for any simple root α. Item (b) and (c) are also satisfied by Lemmas 3.18, 3.19 and Corollary 3.20.

Therefore, we can deform $\bar{\rho}$ to a continuous representation $\rho : \Gamma_{\mathbb{Q}} \to G(\mathcal{O})$ satisfying the prescribed local conditions on S and the Ramakrishna condition on a set of auxiliary primes disjoint from S. We write G_ρ for the Zariski closure of the image of ρ in $G(\bar{\mathbb{Q}}_l)$. By [Patrikis 2016, Lemma 7.7] G_ρ is reductive. By Proposition 4.16, the Steinberg condition at p ensures that G_ρ contains a regular unipotent element (see the proof of [Patrikis 2016, Theorem 8.4]). By a theorem of Dynkin [Saxl and Seitz 1997, Theorem A], G_ρ is then of type A_1 or A_2 for $G = \mathrm{SL}_3$, type A_1, G_2 or B_3 for $G = \mathrm{Spin}_7$, and type A_1, F_4, or E_6 for $G = E_6^{\mathrm{sc}}$. But $\alpha(\rho(\mathrm{Fr}_r))$, $\alpha \in \Delta$ are distinct, so $G_\rho = G(\bar{\mathbb{Q}}_l)$ in all three cases (see the proof of [Patrikis 2016, Lemma 7.8]). □

Remark 4.19. As we have flexibilities in choosing $g \in G(\mathcal{O})$ lifting $\bar{\rho}(\mathrm{Fr}_r)$, it is easy to see that there are infinitely many lifts ρ that are nonconjugate in G^{ad}.

4B3. SL_2. The alternating group A_n admits a unique nontrivial central extension $\tilde{A}_n$ by $\mathbb{Z}/2\mathbb{Z}$ for $n \neq 6, 7$. By a result of N. Vila and J.-F. Mestre (which was proven independently, see [Serre 1992]), $\tilde{A}_n$ can be realized as a Galois group over $\mathbb{Q}$. In particular, we get a surjection $\bar{r} : \Gamma_{\mathbb{Q}} \twoheadrightarrow \tilde{A}_5$. On the other hand, $\tilde{A}_5$ can be described as follows: the symmetries of an icosahedron induce a 3-dimensional irreducible faithful representation of A_5, i.e., there is an injective homomorphism $A_5 \to \mathrm{SO}(3)$. The pullback of A_5 along the two-fold covering map $\mathrm{SU}(2) \twoheadrightarrow \mathrm{SO}(3)$ is a nontrivial central extension of A_5 by $\mathbb{Z}/2\mathbb{Z}$, hence is isomorphic to $\tilde{A}_5$. In particular, we get an embedding $\tilde{A}_5 \to \mathrm{SL}_2(\mathbb{C})$. As the matrix entries of the image lie in a finite extension of $\mathbb{Q}$, we can choose a finite extension k of $\mathbb{F}_l$ for which there is an embedding $\tilde{A}_5 \to \mathrm{SL}_2(k)$. Precomposing it with $\bar{r}$, we obtain a representation $\Gamma_{\mathbb{Q}} \to \mathrm{SL}_2(k)$ which we denote by $\bar{\rho}$. It is easy to see that the adjoint module $\bar{\rho}(\mathfrak{sl}_2(k))$ is irreducible.

Let S be a finite set of places containing the archimedean place and all places where $\bar{\rho}$ is ramified. We impose no condition at l, and take the minimal prime to l deformation condition at all other places in S, which is legitimate since the residual image has order prime to l for $l > 120$. Let $\Sigma \in \tilde{A}_5$ be an element of order 4, whose image in $\mathrm{SL}_2(k)$ is conjugate to $\mathrm{diag}(\sqrt{-1}, -\sqrt{-1})$. As $\tilde{A}_5$ has trivial abelian quotient, $\mathbb{Q}(\mu_l)$ and $\mathbb{Q}(\bar{\rho})$ are linearly disjoint over $\mathbb{Q}$. So there is an element $\sigma \in \Gamma_{\mathbb{Q}}$ such that $\bar{\rho}(\sigma) = \Sigma$ and $\bar{\kappa}(\sigma) = -1$. By Chebotarev's density theorem, there is a prime $p \notin S$ for which $\mathrm{Fr}_p = \sigma$. Therefore, $\bar{\rho}|_{\Gamma_p}$ is of Steinberg type and we take the Steinberg deformation condition at p.

Proposition 4.20. *For $\bar{\rho} : \Gamma_{\mathbb{Q}} \to \mathrm{SL}_2(k)$ defined as above and for almost all primes l, there is an l-adic lift*

$$\rho : \Gamma_{\mathbb{Q}} \to \mathrm{SL}_2(\mathcal{O})$$

of $\bar{\rho}$ with Zariski-dense image in $\mathrm{SL}_2(\bar{\mathbb{Q}}_l)$.

Proof. We first show that Theorem 3.4 applies to $\bar{\rho}$. Item (1) is satisfied: the left hand side of the inequality equals the right hand side. Item (2) is satisfied since $|\mathrm{Gal}\,(K/\mathbb{Q})|$ has order prime to l by the definition of $\bar{\rho}$. Item (3) is satisfied since $\bar{\rho}(\mathfrak{g})$ and $\bar{\rho}(\mathfrak{g})(1)$ are nonisomorphic. For item (4), we take σ to be as above, the connected component of whose centralizer is denoted T, and take α to be a root of $\Phi(G, T)$. So (a) is satisfied. As $\bar{\rho}(\mathfrak{g})$ is irreducible, (b) and (c) are satisfied.

Therefore, we can deform $\bar{\rho}$ to a continuous representation $\rho : \Gamma_{\mathbb{Q}} \to \mathrm{SL}_2(\mathcal{O})$ satisfying the prescribed local conditions on S and the Ramakrishna condition on a set of auxiliary primes disjoint from S. We write G_ρ for the Zariski closure of the image of ρ in $\mathrm{SL}_2(\bar{\mathbb{Q}}_l)$. As $\bar{\rho}(\mathfrak{g})$ is irreducible, $\mathrm{Lie}(G_\rho)$ is either trivial or $\mathfrak{g}_{\bar{\mathbb{Q}}_l}$. If the former were true, then G_ρ would be finite. But $\rho|_{\Gamma_p}$ is Steinberg, so in particular the image of ρ is infinite, a contradiction. Thus $\mathrm{Lie}(G_\rho) = \mathfrak{g}_{\bar{\mathbb{Q}}_l}$. □

Now we finish the proof of Theorem 1.2 with the congruence conditions removed. For G a simple but not simply connected group, suppose there is a homomorphism $\rho_l : \Gamma_{\mathbb{Q}} \to G^{\mathrm{sc}}(\bar{\mathbb{Q}}_l)$ with Zariski-dense image. We compose ρ_l with the covering projection $G^{\mathrm{sc}}(\bar{\mathbb{Q}}_l) \twoheadrightarrow G(\bar{\mathbb{Q}}_l)$, the resulting map has Zariski-dense image in $G(\bar{\mathbb{Q}}_l)$. Propositions 4.14, 4.18 and 4.20 prove the cases of a simple, simply connected classical group, E_6 and E_7. On the other hand, the remaining cases G_2, F_4, and E_8 have already been established in [Patrikis 2016] in a way similar to the proof of Proposition 4.18. So Theorem 1.2 is proved.

In order to remove the congruence conditions for G of type B_n, C_n and E_7, we impose the same local deformation conditions as specified in the paragraph preceding Lemma 4.11, but then use Theorem 3.21 instead of Theorem 3.16. By Lemmas 4.11 and 3.22, the assumptions in Theorem 3.21 are all met. Therefore, we obtain a characteristic zero lift of $\bar{\rho}$ satisfying the prescribed local conditions for *all* large enough primes l. Then the proof of Proposition 4.14 shows that the lift has full monodromy group for all large enough primes.

5. Connected reductive groups as monodromy groups

Following [Milne 2007], a connected algebraic group G is said to be an *almost-direct product* of its algebraic subgroups $G_1, \ldots, G_n$ if the map

$$G_1 \times \cdots \times G_n \to G : (g_1, \ldots, g_n) \mapsto g_1 \cdots g_n$$

is a surjective homomorphism with finite kernel; in particular, this means that the G_i's commute with each other and each G_i is normal in G.

The following proposition is [Milne 2007, Corollary 4.4].

Proposition 5.1. *An algebraic group is semisimple if and only if it is an almost direct product of simple algebraic groups. (Here a simple algebraic group is called almost simple in [Milne 2007]).*

Proposition 5.2 (Goursat's lemma). *Let G_1, G_2 be groups, let H be a subgroup of $G_1 \times G_2$ such that the two projections $p_1 : H \to G_1$, $p_2 : H \to G_2$ are surjective. Let N_1 and N_2 be the kernels of p_2 and p_1, respectively. Then the image of H in $G_1/N_1 \times G_2/N_2$ is the graph of an isomorphism $G_1/N_1 \cong G_2/N_2$.*

Proposition 5.3. *Let G be a connected semisimple group, then there are continuous homomorphisms*

$$\Gamma_{\mathbb{Q}} \to G(\bar{\mathbb{Q}}_l)$$

with Zariski-dense image for large enough l.

Proof. This will follow from Goursat's lemma, Theorem 1.2 and Remark 1.3. By Proposition 5.1, it suffices to prove the case when G is the direct product of simply connected simple algebraic groups. We may decompose G into "isotypic factors": $G = G_1 \times \cdots \times G_n$, where G_i is the direct product of copies of some simple algebraic group, and G_i, G_j have different types for $i \neq j$. Suppose we are given $\rho_i : \Gamma_{\mathbb{Q}} \to G_i(\bar{\mathbb{Q}}_l)$ with Zariski-dense image for each i and let $\rho := (\rho_1, \ldots, \rho_n) : \Gamma_{\mathbb{Q}} \to G(\bar{\mathbb{Q}}_l)$, whose Zariski-closure is denoted by H. Then H is an algebraic subgroup of $G(\bar{\mathbb{Q}}_l)$ for which $\mathrm{pr}_i(H) = G_i(\bar{\mathbb{Q}}_l)$. Since $G_i(\bar{\mathbb{Q}}_l)$ and $G_j(\bar{\mathbb{Q}}_l)$ share no common nontrivial quotients for $i \neq j$, Proposition 5.2 implies that $H = G(\bar{\mathbb{Q}}_l)$.

It remains to prove the case when G is a direct product of copies of some simply connected simple algebraic group. Write $G = K^n$ with K simple. We first assume $K \neq \mathrm{SL}_2$. By Section 4B1 and 4B2 (especially Remarks 4.15 and 4.19), there exists a prime p and a homomorphism $\rho_i : \Gamma_{\mathbb{Q}} \to K(\bar{\mathbb{Q}}_l)$ for $1 \leq i \leq n$ such that ρ_i has Zariski-dense image and is unramified at p with $\rho_i(\mathrm{Fr}_p)$ a regular semisimple element in $K(\bar{\mathbb{Q}}_l)$, and for $i \neq j$, the images of $\rho_i(\mathrm{Fr}_p)$ and $\rho_j(\mathrm{Fr}_p)$ in $K^{\mathrm{ad}}(\bar{\mathbb{Q}}_l)$ are not conjugate by an automorphism of K^{ad}. Now we use Proposition 5.2 and induction on n to show that $\rho := \prod_i \rho_i : \Gamma_{\mathbb{Q}} \to G(\bar{\mathbb{Q}}_l)$ has Zariski-dense image. This is clear when $n = 1$. Suppose this is true for $n-1$, so that $\prod_{i<n} \rho_i : \Gamma_{\mathbb{Q}} \to K^{n-1}(\bar{\mathbb{Q}}_l)$ has Zariski-dense image. Let H be the Zariski-closure of the image of ρ and apply Proposition 5.2 to $p_1 : H \to G_1 = K^{n-1}$, $p_2 : H \to G_2 = K$, which are surjective by assumption, we see that the image of H in $G_1/N_1 \times G_2/N_2$ is the graph of an isomorphism $G_1/N_1 \cong G_2/N_2$. But since $G_2 = K$ is a simple, simply connected algebraic group, $N_2 = K$ or $N_2 \subset Z(K)$. Also note that N_1 is either K^{n-1} or isogenous to the product of $n-2$ factors in K^{n-1}. If $N_2 = K$ and $N_1 = K^{n-1}$, H must be G; otherwise, $N_2 \subset Z(K)$ and N_1 is isogenous to the product of $n-2$ factors in K^{n-1}, so the isomorphism $G_1/N_1 \times G_2/N_2$ induces an isomorphism between two factors in $G^{\mathrm{ad}} = (K^{\mathrm{ad}})^n$. But this is impossible, as for any $i \neq j$, the images of $\rho_i(\mathrm{Fr}_p)$ and $\rho_j(\mathrm{Fr}_p)$ in $K^{\mathrm{ad}}(\bar{\mathbb{Q}}_l)$ are not conjugate by an automorphism of K^{ad}. Therefore, $H = G$.

Now let $K = \mathrm{SL}_2$. By the argument in the previous paragraph, it suffices to show that for any n, we can construct homomorphisms $\rho_i : \Gamma_{\mathbb{Q}} \to \mathrm{SL}_2(\bar{\mathbb{Q}}_l)$, $1 \leq i \leq n$ such that ρ_i has Zariski-dense image and for

$i \neq j$, ρ_i^{ad} and ρ_j^{ad} are not conjugated by an automorphism of PGL_2. By the construction in [Serre 1992, 9.3] there are infinitely many homomorphisms $r : \Gamma_{\mathbb{Q}} \twoheadrightarrow \tilde{A}_5$ such that the composites $\Gamma_{\mathbb{Q}} \xrightarrow{r} \tilde{A}_5 \to A_5$ are ramified at different sets of finite primes. By Section 4B3 we then obtain infinitely many homomorphisms $\bar{\rho} : \Gamma_{\mathbb{Q}} \to \mathrm{SL}_2(k)$ such that the corresponding homomorphisms $\bar{\rho}^{\mathrm{ad}} : \Gamma_{\mathbb{Q}} \to \mathrm{PGL}_2(k)$ are ramified at different sets of finite primes. By Proposition 4.20, we can deform $\bar{\rho}$ to characteristic zero with Zariski-dense image. We take n of them, denoted by $\rho_1, \ldots, \rho_n : \Gamma_{\mathbb{Q}} \to \mathrm{SL}_2(\overline{\mathbb{Q}}_l)$. Suppose for some i, j with $1 \leq i \neq j \leq n$, ρ_i^{ad} and ρ_j^{ad} were conjugate by an automorphism of PGL_2, then their mod l reductions $\bar{\rho}_i^{\mathrm{ad}}$ and $\bar{\rho}_j^{\mathrm{ad}}$ would be conjugate as well. But since there exists a prime p for which $\bar{\rho}_i^{\mathrm{ad}}$ is ramified at p and $\bar{\rho}_j^{\mathrm{ad}}$ is unramified at p, $\bar{\rho}_i^{\mathrm{ad}}$ and $\bar{\rho}_j^{\mathrm{ad}}$ cannot be conjugate, a contradiction. □

Lemma 5.4. *Let n be a positive integer and $T = (\mathbb{G}_m)^n$. Then there is a continuous map $\iota : \mathbb{Z}_l \to T(\mathbb{Q}_l)$ with Zariski-dense image.*

Proof. The group T has only countably many connected proper Zariski-closed subgroups, so one can pick a line L in $\mathrm{Lie}(T)_{\mathbb{Q}_l}$ avoiding the tangent spaces to all such proper subgroups (since $\mathbb{Q}_l$ is uncountable). A small compact neighborhood of 0 in L exponentiates to a compact subgroup C of $T(\mathbb{Q}_l)$ whose Zariski closure has identity component that cannot be a proper algebraic subgroup of T, so C is Zariski-dense in T. □

Now we can prove Theorem 1.1: Let G be a connected reductive group, then G is a quotient of the product of G^{der} (a semisimple group) and $Z(G)^0$ (a torus). Proposition 5.3 and Lemma 5.4 then allow us to build a homomorphism from $\Gamma_{\mathbb{Q}}$ to $G(\overline{\mathbb{Q}}_l)$ with Zariski-dense image for large enough l.

Acknowledgement

I am extremely grateful for the patience, generosity, and support of my advisor Stefan Patrikis, without whom this work would be impossible. I would also like to thank Gordan Savin for his patience and generous support on a few projects I worked on earlier in my graduate career. I thank Jeffrey Adams for answering my questions on his recent joint work with Xuhua He. The University of Utah mathematics department was a supportive and intellectually enriching environment, and I am grateful to all of the faculty members, postdocs, and graduate students who made it so. I would also like to thank the anonymous referees for correcting the mistakes and making many useful comments on an earlier draft of this paper.

References

[Adams and He 2017] J. Adams and X. He, "Lifting of elements of Weyl groups", *J. Algebra* **485** (2017), 142–165. MR Zbl

[Bourbaki 1968] N. Bourbaki, *Groupes et algèbres de Lie, Chapitres IV–VI*, Actualités Scientifiques et Industrielles **1337**, Hermann, Paris, 1968. Translated as *Lie groups and Lie algebras, chapters 4–6*, Springer, 2002. MR Zbl

[Buzzard and Gee 2014] K. Buzzard and T. Gee, "The conjectural connections between automorphic representations and Galois representations", pp. 135–187 in *Automorphic forms and Galois representations, I* (Durham, 2011), edited by F. Diamond et al., London Math. Soc. Lecture Note Ser. **414**, Cambridge Univ. Press, 2014. MR Zbl

[Clozel et al. 2008] L. Clozel, M. Harris, and R. Taylor, "Automorphy for some l-adic lifts of automorphic mod l Galois representations", *Publ. Math. Inst. Hautes Études Sci.* **108** (2008), 1–181. MR Zbl

[Cornut and Ray 2018] C. Cornut and J. Ray, "Generators of the pro-p Iwahori and Galois representations", *Int. J. Number Theory* **14**:1 (2018), 37–53. MR Zbl

[Dettweiler and Reiter 2010] M. Dettweiler and S. Reiter, "Rigid local systems and motives of type G_2", *Compos. Math.* **146**:4 (2010), 929–963. MR Zbl

[Fakhruddin et al. 2018] N. Fakhruddin, C. Khare, and S. Patrikis, "Lifting irreducible Galois representations", preprint, 2018. arXiv

[Fontaine and Mazur 1995] J.-M. Fontaine and B. Mazur, "Geometric Galois representations", pp. 41–78 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

[Fontaine and Messing 1987] J.-M. Fontaine and W. Messing, "p-adic periods and p-adic étale cohomology", pp. 179–207 in *Current trends in arithmetical algebraic geometry* (Arcata, CA, 1985), edited by K. A. Ribet, Contemp. Math. **67**, Amer. Math. Soc., Providence, RI, 1987. MR Zbl

[Gan and Yu 2003] W. T. Gan and J.-K. Yu, "Schémas en groupes et immeubles des groupes exceptionnels sur un corps local, I: Le groupe G_2", *Bull. Soc. Math. France* **131**:3 (2003), 307–358. MR Zbl

[Greenberg 2016] R. Greenberg, "Galois representations with open image", *Ann. Math. Qué.* **40**:1 (2016), 83–119. MR Zbl

[Humphreys 1995] J. E. Humphreys, *Conjugacy classes in semisimple algebraic groups*, Math. Surveys and Monographs **43**, Amer. Math. Sci., Providence, RI, 1995. MR Zbl

[Katz 2018] N. M. Katz, "A note on Galois representations with big image", preprint, 2018, Available at https://tinyurl.com/katzbig.

[Kostant 1959] B. Kostant, "The principal three-dimensional subgroup and the Betti numbers of a complex simple Lie group", *Amer. J. Math.* **81** (1959), 973–1032. MR Zbl

[Kret and Shin 2016] A. Kret and S. W. Shin, "Galois representations for general symplectic groups", preprint, 2016. arXiv

[LMFDB 2013] LMFDB Collaboration, "The L-functions and modular forms database", 2013, Available at http://www.lmfdb.org.

[Milne 2007] J. S. Milne, "Semisimple algebraic groups in characteristic zero", preprint, 2007. arXiv

[Neukirch et al. 2000] J. Neukirch, A. Schmidt, and K. Wingberg, *Cohomology of number fields*, Grundlehren der Math. Wissenschaften **323**, Springer, 2000. MR Zbl

[Patrikis 2016] S. Patrikis, "Deformations of Galois representations and exceptional monodromy", *Invent. Math.* **205**:2 (2016), 269–336. MR Zbl

[Ramakrishna 2002] R. Ramakrishna, "Deforming Galois representations and the conjectures of Serre and Fontaine–Mazur", *Ann. of Math.* (2) **156**:1 (2002), 115–154. MR Zbl

[Ribet 1985] K. A. Ribet, "On l-adic representations attached to modular forms, II", *Glasgow Math. J.* **27** (1985), 185–194. MR Zbl

[Saxl and Seitz 1997] J. Saxl and G. M. Seitz, "Subgroups of algebraic groups containing regular unipotent elements", *J. London Math. Soc.* (2) **55**:2 (1997), 370–386. MR Zbl

[Serre 1992] J.-P. Serre, *Topics in Galois theory*, Res. Notes in Math. **1**, Jones and Bartlett, Boston, MA, 1992. MR Zbl

[Serre 1996] J.-P. Serre, "Exemples de plongements des groupes $\mathrm{PSL}_2(\mathbb{F}_p)$ dans des groupes de Lie simples", *Invent. Math.* **124**:1-3 (1996), 525–562. MR Zbl

[Taylor 2003] R. Taylor, "On icosahedral Artin representations, II", *Amer. J. Math.* **125**:3 (2003), 549–566. MR Zbl

[Weston 2004] T. Weston, "Unobstructed modular deformation problems", *Amer. J. Math.* **126**:6 (2004), 1237–1252. MR Zbl

[Yun 2014] Z. Yun, "Motives with exceptional Galois groups and the inverse Galois problem", *Invent. Math.* **196**:2 (2014), 267–337. MR Zbl

Communicated by Michael Rapoport
Received 2018-06-15 Revised 2018-10-18 Accepted 2018-11-20

tang@math.utah.edu *Department of Mathematics, University of Utah, Salt Lake City, UT, United States*

ALGEBRA AND NUMBER THEORY 13:6 (2019)
dx.doi.org/10.2140/ant.2019.13.1395

Weyl bound for p-power twist of GL(2) L-functions

Ritabrata Munshi and Saurabh Kumar Singh

Let f be a cuspidal eigenform (holomorphic or Maass) for the congruence group $\Gamma_0(N)$ with N square-free. Let p be a prime and let χ be a primitive character of modulus p^{3r}. We shall prove the Weyl-type subconvex bound

$$L\big(\tfrac{1}{2}+it, f\otimes\chi\big) \ll_{f,t,\varepsilon} p^{r+\varepsilon},$$

where $\varepsilon > 0$ is any positive real number.

1. Introduction

Bounding automorphic L-functions on the critical line $\mathrm{Re}(s) = \frac{1}{2}$ is a central problem in the analytic theory of L-functions. The functional equation and the Phragmen–Lindelöf principle from complex analysis yield the convexity bound $L\big(\frac{1}{2}+it,\pi\big) \ll C(\pi,t)^{1/4+\varepsilon}$ where $C(\pi,t)$ is the analytic conductor of the L-function, as defined by Iwaniec and Sarnak (see equation (31) of [Iwaniec and Sarnak 1999]), whereas the grand Riemann hypothesis (GRH) implies the Lindelöf hypothesis which predicts the bound $C(\pi,t)^{\varepsilon}$ for any $\varepsilon > 0$. Any bound with exponent smaller than $\frac{1}{4}$ is called a subconvex bound. In this context the Weyl exponent $\frac{1}{6}$, which is one-third of the way down from convexity towards Lindelöf, is a known barrier which has been achieved only for a handful of families. Indeed the only known "arithmetic case" (level aspect) is given by the fundamental work of Conrey and Iwaniec [2000]. For χ_q the quadratic character modulo q (which is square-free and odd) they established: (i) $L\big(\frac{1}{2}+it,\chi_q\big) \ll_t q^{1/6+\varepsilon}$ with polynomial dependence on t. (ii) $L\big(\frac{1}{2}, f\otimes\chi_q\big) \ll q^{1/3+\varepsilon}$ for f a primitive GL(2) cusp form of level dividing q. Note that while the former result is flexible and applies to any point on the critical line, the latter result only applies at the central point, as the nonnegativity of the L-value plays a central role in their argument. Our main objective here is to establish the following Weyl-type bound.

Theorem 1.1. *Let f be a holomorphic Hecke eigenform, or a Maass cusp form for the congruence subgroup $\Gamma_0(N)$ with N square-free. Let χ be a primitive character of modulus p^{3r} where p is a prime and r is a natural number. We have*

$$L\big(\tfrac{1}{2}+it, f\otimes\chi\big) \ll_{f,t,\varepsilon} p^{r+\varepsilon}.$$

MSC2010: primary 11F66; secondary 11F55, 11M41.
Keywords: Maass forms, Hecke eigenforms, Voronoi summation formula, Poisson summation formula.

This is the first instance where the Weyl exponent is achieved in the level aspect for a family of L-functions which are not self-dual. Our method is soft and does not rely on the Riemann hypothesis for varieties over finite fields. Indeed the character sums that we encounter are treated in an elementary manner. It is quite surprising that our method could yield such a strong bound. But recently it has been established (see [Munshi 2018; Aggarwal and Singh 2017]) that both the Weyl and the Burgess bound, for both GL(1) and GL(2) L-functions, follow from the method developed in the series [Munshi 2014; 2015a; 2015b; 2015c].

Recall that the twisted automorphic L-function of degree two associated to (f, χ) is defined by the Dirichlet series

$$L(s, f\otimes\chi) := \sum_{n=1}^{\infty}\frac{\lambda_f(n)\chi(n)}{n^s},$$

where $\mathrm{Re}(s) > 1$ and $\lambda_f(n)$ are the normalized Fourier coefficients of f. These extend to an entire function and satisfy a functional equation relating $L(s, f\otimes\chi)$ to $L(1-s, \bar{f}\otimes\bar{\chi})$. In the family we are considering the form f and the point $\frac{1}{2}+it$ are fixed, but the character χ varies and $p\to\infty$. The above result will be derived as a special case of the following.

Theorem 1.2. *Let f be a holomorphic Hecke eigenform or a Hecke–Maass cusp form for $\Gamma_0(N)$ with N square-free. Let χ be a primitive character of modulus p^r, where p is a prime. We have*

$$L\big(\tfrac{1}{2}+it, f\otimes\chi\big) \ll_{f,\epsilon} p^{1/2(r-\lfloor r/3\rfloor)+\epsilon},$$

where $\epsilon > 0$ is any positive real number and $\lfloor x\rfloor$ is the greatest integer less than or equal to x.

Let us briefly recall the history of subconvexity bounds for L-functions. We will only focus on the results which are related to our case. The convexity bound for the Riemann zeta function is given by $\zeta\big(\frac{1}{2}+it\big) \ll t^{1/4+\varepsilon}$. For a Dirichlet L-function associated with a primitive Dirichlet character χ of modulus q, the convexity bound is given by $L\big(\frac{1}{2}, \chi\big) \ll q^{1/4+\varepsilon}$. The Lindelöf hypothesis asserts that the exponent $\frac{1}{4}+\varepsilon$ can be replaced by ε. The subconvexity bound for $\zeta(s)$ was first proved by Hardy and Littlewood, based on the work of Weyl [1916]. Establishing a bound for exponential sums, it has been proved that (see also [Titchmarsh 1986, page 99, Theorem 5.5])

$$\zeta\big(\tfrac{1}{2}+it\big) \ll t^{1/6}\log^{3/2} t. \tag{1}$$

It was first written down by Landau [1924] in a slightly refined form, and has been generalized to all Dirichlet L-functions. Since then it has been improved by several people. The best known result with exponent $\frac{13}{84}\approx 0.15476$ is due to Bourgain [2017]. On the other hand, the q-aspect subconvexity bound was first proved by Burgess [1963]. Using an ingenious technique of completing short character sums and utilizing the Riemann hypothesis for curves over finite fields (Weil's result), he proved that

$$L(s, \chi) \ll_\epsilon q^{3/16+\varepsilon}, \tag{2}$$

for fixed s with $\mathrm{Re}\, s = \frac{1}{2}$ and for any $\varepsilon > 0$. Heath-Brown [1978] proved the hybrid bound (bound in both parameters q and t together) for Dirichlet L-functions. The Burgess exponent $\frac{3}{16}$, which is one-fourth of the way down from convexity towards the Lindelöf, has come to be realized as a natural barrier. However, stronger bounds can be shown under suitable factorization hypothesis on the modulus. Indeed, in principle the family $L\left(\frac{1}{2}, \chi\right)$ where χ runs over characters modulo p^r with $r \to \infty$ should behave like the family $\zeta\left(\frac{1}{2}+it\right)$ with $t \to \infty$. Only recently a suitable p-adic analogue of the van der Corput method has been introduced by Milićević [2016], and he has been able to obtain a sub-Weyl exponent 0.1645 for this family. More precisely, for χ primitive Dirichlet character modulo $q = p^n$ he proved that for any given $\theta > \theta_0 \approx 0.1645$, there is a $j \geq 0$ such that

$$L\left(\tfrac{1}{2}, \chi\right) \ll p^j q^{\theta} (\log q)^{1/2}.$$

So it follows that we have a subconvexity exponent which is less than $\frac{1}{6}$ for a prime power modulus $q = p^n$ with $n \geq n_0$, a sufficiently large number.

The t-aspect Weyl exponent for GL(2) L-functions was first proved by Good [1982], for holomorphic modular forms, using the spectral theory of automorphic functions. Jutila [1987] has given an alternative proof, based only on the functional properties of $L(f, s)$ and $L(s, f \otimes \psi)$, where ψ is an additive character. The arguments used in his proof were flexible enough to be adopted for the Maass cusp forms, as was shown by Meurman [1990]. However, the character twist aspect subconvexity bound required some more new ideas. It was first obtained by Duke, Friedlander and Iwaniec using a new form of the circle method and the amplification technique. Assuming χ to be a primitive character of modulus q and $\operatorname{Re} s = \frac{1}{2}$, they obtained (see [Duke et al. 1993, Theorem 1])

$$L(s, f \otimes \chi) \ll_f |s|^2 q^{5/11} \tau^2(q) \log q,$$

where $\tau(q)$ is the divisor function. In the case of a general holomorphic cusp form, Bykovskii [1996] used a trace formula expressing the mean values of cusp form (see [Bykovskii 1996, page 1, line 1]) to derive the Burgess exponent $\frac{3}{8}$. In the case of Maass form subconvexity bound obtained by Harcos. Refining the arguments used in [Bykovskii 1996], Blomer and Harcos [2008] obtained the Burgess exponent $\frac{3}{8}$ for a more general holomorphic or Maass cusp form. To date, the Weyl exponent $\frac{1}{3}$ has only been achieved for quadratic characters, courtesy of the fundamental work of Conrey and Iwaniec [2000], as we have already mentioned above. Extending the above mentioned result of Milićević to GL(2) L-functions, Blomer and Milićević [2015, Theorem 2] obtained

$$L\left(\tfrac{1}{2}+it, f \otimes \chi\right) \ll_{f,\varepsilon} (1+|t|)^{5/2} p^{7/6} q^{1/3+\varepsilon},$$

where f is a holomorphic or Maass newform for $\mathrm{SL}(2, \mathbb{Z})$, and χ is a primitive character of conductor $q = p^n$, with p an odd prime. This yields a subconvexity bound for $n > 7$ and improves the Burgess exponent as soon as $n > 27$, and the exponent tends to Weyl exponent as $n \to \infty$. Though there is sufficient room for improvements in the above estimate (as the authors themselves comment in [Blomer and Milićević 2015]), it is inherent in their method that the Weyl exponent can never be achieved for any given n. In this paper we propose a different approach which produces an improvement over the known bounds. In Theorem 1.2, we are able to provide an improvement on the Burgess exponent as soon as $n \geq 3$, with the exception when $n = 4, 8$ (in this case our exponent is same as Burgess exponent) and $n = 5$. Of course the most interesting outcome of our result is that we are able to achieve the Weyl exponent when $n \geq 3$ and $n \equiv 0 \pmod 3$. In the next section we briefly explain the method of the proof, which is a rendition of [Munshi 2015a].

2. Sketch of the proof

We start with the approximate functional equation as given in [Iwaniec and Kowalski 2004, Proposition 5.4]. Taking a smooth dyadic subdivision we arrive at

$$L\left(\tfrac{1}{2}+it, f\otimes\chi\right) \ll_{f,A} N^{\varepsilon} \sup_{N\leq P^{1+\varepsilon}} \frac{|S(N)|}{N^{1/2}} + P^{-A},$$

where $P = p^r$ is the modulus of the character and $S(N)$ is a dyadic sum which is given by

$$S(N) := \sum_{n=1}^{\infty} \lambda_f(n)\chi(n)V\left(\frac{n}{N}\right), \tag{3}$$

where $V(x)$ is a smooth bump function supported on the interval $[1, 2]$ and satisfies $V^{(j)}(x) \ll_j (1+|t|)^j$. As the implied constant in Theorem 1.2 is allowed to depend on t, we can and will from now on assume that $t = 0$. Our method is not sensitive to small perturbations like this. A careful study of the proof shows that the eventual implied constant grows at most polynomially with t. Now trivially estimating, we obtain $S(N) \ll N^{1+\varepsilon}$. We shall examine $S(N)$ in following steps. For simplicity let us focus on the case $r \equiv 0 \pmod 3$.

Step 1 (applying circle method). We shall apply Kloosterman's version of the circle method (see Lemma 3.3) with the conductor lowering mechanism introduced by the first author in [Munshi 2015a]. We obtain the sum

$$S(N) = 2\,\mathrm{Re}\int_0^1 \sum_{1\leq q\leq Q}\sideset{}{^\star}\sum_{Q<q\leq q+Q} \frac{1}{aqp^{\ell}} \sum_{b(p^{\ell})} \left\{\sum_{n\sim N} \lambda_f(n) e\left(\frac{(\bar{a}+bq)n}{p^{\ell}q}\right)\right\}\left\{\sum_{m\sim N} \chi(m) e\left(-\frac{(\bar{a}+bq)m}{p^{\ell}q}\right)\right\} dx,$$

where Q is taken to be $Q = N^{1/2}/p^{\ell/2}$. Trivially estimating after first step, we have $S(N) \ll N^{2+\varepsilon}$.

Step 2 (applying Poisson summation formula). In this step we shall apply the Poisson summation formula to the sum over m. The character χ has conductor p^r and the additive conductor has a size q. Hence the total conductor for the sum over m has size $p^r q$. So the dual sum is essentially supported up to the size qp^r/N. We observe that after application of the Poisson formula we are able to save $N/\sqrt{p^r q}$ from the sum over m. Evaluating the character sum, we also observe that $Q < a \leq Q+q$ can be determined uniquely by the congruence relation $a \equiv \bar{m}\, p^{r-\ell} \bmod q$. In particular a does not depend on n. The total saving after the first step is given by

$$\frac{N}{\sqrt{p^r q}} \times \sqrt{q} = \frac{N}{p^{r/2}}.$$

Trivially estimating after the second step we obtain $S(N) \ll Np^{r/2}$.

Step 3 (applying Voronoi summation formula). We shall now apply the Voronoi summation formula to the sum over n, which has conductor of size $p^\ell q$. The dual length is essentially supported up to the size $p^{2\ell}q^2/N$. We are able to save $N/p^\ell q$ from the Voronoi summation formula and $p^{\ell/2}$ by assuming square root cancellation in exponential sum over b. Total saving in the third step is

$$\frac{N}{p^\ell q} \times p^{\ell/2} = \frac{N}{p^{\ell/2} q}.$$

Trivially estimating after the third step, we observe that $S(N) \ll N^{1/2} p^{r/2+\varepsilon}$, which shows that we are on the boundary. We are left with a sum of the form:

$$S(N) = \sum_{n \ll p^\ell N^\varepsilon} \lambda_f(n) \Bigg[\sum_{1 \le q \le Q} \sum_{m \ll \frac{N^\varepsilon Q p^r}{N}} \frac{\chi(q)}{aq^2} \mathcal{C}(n,m,q) e\bigg(-n \frac{p^r \overline{p^{2\ell}} \overline{m}}{q} \bigg) \mathcal{I}(x,q,m) \mathcal{J}(x,n,q) \Bigg],$$

where the character sum is given by

$$\mathcal{C}(n,m,q) = \sum_{b \bmod p^\ell}^{\star} \bar{\chi}(m - (\bar{a} + bq) p^{r-\ell}) e\bigg(-n \frac{\overline{a + bq\bar{q}}}{p^\ell} \bigg)$$

and the function $\mathcal{J}(x,n,q)$ is of size $O(1)$, and is not oscillatory with respect to n.

Step 4 (Cauchy–Schwarz inequality and Poisson summation formula). To obtain additional savings, we apply the Cauchy–Schwarz inequality to get rid of the Fourier coefficients. But this process also squares the amount we need to save. We now open the absolute square and then interchange the summation over n. Applying the Poisson summation formula we are able to save $Q^2 p^r/N = p^{r-\ell}$ from the diagonal terms and $p^{\ell/2}$ from the nondiagonal terms. We observe that optimal choice for ℓ is given by $\ell = 2r/3$. Substituting the value of ℓ we obtain

$$S(N) \ll \frac{N^{1/2} p^{r/2+\varepsilon}}{p^{\ell/4}} \ll N^{1/2} p^{r/3+\varepsilon}.$$

Upon substituting this bound in the bound we obtained from the approximate functional equation the Weyl bound follows. Observe that when r is not divisible by 3 then we are not allowed to pick the above optimal choice for ℓ, and we have to choose the best possible which is $[2r/3]$. In the following sections we shall provide the proof of the theorem in detail.

3. Preliminaries

To keep the notations simple we will focus on the case of full level. Our argument is robust and is not sensitive to the nature of the fixed form f. We will present our argument in detail for Maass forms of full level. The case of Maass forms is traditionally considered harder. The reader will have no problem to see how the arguments can be adopted in the case of general square-free level with general nebentypus. In principle, our method should work even for levels which are not square-free, but we refrain from including that due to the lack of a suitable Voronoi summation formula. In this section we recall some basic facts about $\mathrm{SL}(2,\mathbb{Z})$ automorphic forms (for details see [Iwaniec 1997; Iwaniec and Kowalski 2004]). Our requirement is minimal, in fact the Voronoi summation formula and the Rankin–Selberg bound (see Lemma 3.2) are all that we shall be using.

Maass cusp forms. Let f be a weight zero Hecke–Maass cusp form with Laplace eigenvalue $\frac{1}{4}+\nu^2$. The Fourier series expansion of f at ∞ is given by

$$f(z)=\sqrt{y}\sum_{n\neq 0}\lambda_f(n)K_{i\nu}(2\pi|n|y)e(nx),$$

where $\lambda_f(1)=1$. Let χ be a primitive Dirichlet character of modulus P. The twisted L-function is defined by

$$L(s,f\otimes\chi):=\sum_{n=1}^{\infty}\frac{\lambda_f(n)\chi(n)}{n^s}$$

for $\mathrm{Re}(s)>1$. It extends to an entire function and satisfies the functional equation

$$\Lambda(s,f\otimes\chi)=\varepsilon(f\otimes\chi)\Lambda(1-s,f\otimes\chi),$$

where $|\varepsilon(f\otimes\chi)|=1$ and

$$\Lambda(s,f\otimes\chi)=\left(\frac{P}{\pi}\right)^s\Gamma\left(\frac{s+i\nu}{2}\right)\Gamma\left(\frac{s-i\nu}{2}\right)L(s,f\otimes\chi).$$

From the functional equation and Phragmen–Lindelöf principle one can derive the convexity bound

$$L\big(\tfrac{1}{2}+it,f\otimes\chi\big)\ll_{f,t,\epsilon}P^{1/2+\epsilon}.$$

We shall require the following Voronoi summation formula for the Maass form. This was first established by T. Meurman [1988] for full level (for general case see appendix A.4 of [Kowalski et al. 2002]).

Lemma 3.1 (Vornoi summation formula). *Let h be a compactly supported smooth function in the interval $(0,\infty)$. Let $\lambda_f(n)$ be the Fourier coefficient of a weight zero Maass form for the full modular group $\mathrm{SL}(2,\mathbb{Z})$, and a, q be positive integers with $(a,q)=1$. We have*

$$\sum_{n=1}^{\infty}\lambda_f(n)e_q(an)h(n)=\frac{1}{q}\sum_{\pm}\sum_{n=1}^{\infty}\lambda_f(\mp n)e_q(\pm\bar{a}n)H^{\pm}\left(\frac{n}{q^2}\right),\tag{4}$$

where $a\bar{a}\equiv 1\pmod{q}$, and

$$H^{-}(y)=\frac{-\pi}{\cosh(\pi\nu)}\int_0^{\infty}h(x)\{Y_{2i\nu}+Y_{-2i\nu}\}(4\pi\sqrt{xy})\,dx,$$
$$H^{+}(y)=4\cosh(\pi\nu)\int_0^{\infty}h(x)K_{2i\nu}(4\pi\sqrt{xy})\,dx,$$

where $Y_{2i\nu}$ and $K_{2i\nu}$ are Bessel functions of first and second kind and $e_q(x)=e^{2\pi ix/q}$.

Remark. When h is supported on the interval $[X, 2X]$ and satisfies $x^j h^{(j)}(x) \ll 1$, then integrating by parts and using the properties of Bessel's function, it is easy to see that the sums on the right hand side of (4) are essentially supported on $n \ll_{f,\varepsilon} q^2(qX)^\varepsilon/X$. For smaller values of n we will use the trivial bound, $H^\pm(n/q^2) \ll X$.

Some lemmas. We first recall the Rankin–Selberg bound for Fourier coefficients.

Lemma 3.2. *Let $\lambda_f(n)$ be the normalized Fourier coefficients of a holomorphic cusp form or of a Maass form. Then for any real number $x \geq 1$, we have*

$$\sum_{1\leq n\leq x} |\lambda_f(n)|^2 \ll_{f,\varepsilon} x^{1+\varepsilon}.$$

Let $\delta : \mathbb{Z} \to \{0, 1\}$ be the Kronecker delta function, which is given by

$$\delta(n) = \begin{cases} 1 & \text{if } n = 0, \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

We have the following lemma, which gives the Fourier–Kloosterman expansion of $\delta(n)$ (see [Iwaniec and Kowalski 2004, page 470, Proposition 20.7]).

Lemma 3.3. *Let $Q \geq 1$ be a real number. We have*

$$\delta(n) = 2\,\mathrm{Re} \int_0^1 \sum_{1\leq q\leq Q<a\leq q+Q} \sideset{}{^\star}\sum \frac{1}{aq} e\left(\frac{n\bar{a}}{q} - \frac{nx}{aq}\right), \tag{6}$$

where $\star$ restricts the summation by $(a, q) = 1$ and $a\bar{a} \equiv 1 \pmod{q}$.

We also need to estimate the exponential integral of the form

$$\mathfrak{I} = \int_a^b g(x) e(f(x))\, dx, \tag{7}$$

where f and g are smooth real valued function on the interval $[a, b]$. Suppose we have $|f'(x)| \geq B$, $|f^{(j)}(x)| \leq B^{1+\varepsilon}$ for $j \geq 2$ and $|g^{(j)}(x)| \ll_j 1$ on the interval $[a, b]$. Then by making the change of variable

$$u = f(x), \qquad f'(x)\, dx = du,$$

we have

$$\mathfrak{I} = \int_{f(a)}^{f(b)} \frac{g(x)}{f'(x)} e(u)\, du \quad (x = f^{-1}(u)).$$

By applying integration by parts, differentiating $g(x)/f'(x)$ j-times and integrating $e(u)$, we have

$$\mathfrak{I} \ll_{j,\varepsilon} B^{-j+\varepsilon}. \tag{8}$$

This will be used at several places to show that certain exponential integrals are negligibly small in the absence of a stationary phase point. Next we consider the case of stationary phase point(i.e., point where derivative vanishes).

Lemma 3.4. *Suppose f and g are smooth real valued functions on the interval $[a,b]$ satisfying*

$$f^{(i)} \ll \frac{\Theta_f}{\Omega_f^i}, \quad g^{(j)} \ll \frac{1}{\Omega_g^j} \quad \text{and} \quad f^{(2)} \gg \frac{\Theta_f}{\Omega_f^2}, \tag{9}$$

for $i = 1, 2$ and $j = 0, 1, 2$. Suppose that $g(a) = g(b) = 0$.

(1) *Suppose f' and f'' do not vanish on the interval $[a,b]$. Let $\Lambda = \min_{x\in[a,b]}|f'(x)|$. Then we have*

$$\mathfrak{I} \ll \frac{\Theta_f}{\Omega_f^2 \Lambda^3}\left(1 + \frac{\Omega_f}{\Omega_g} + \frac{\Omega_f^2}{\Omega_g^2}\frac{\Lambda}{\Theta_f/\Omega_f}\right).$$

(2) *Suppose that $f'(x)$ changes sign from negative to positive at $x = x_0$ with $a < x_0 < b$. Let $\kappa = \min\{b - x_0, x_0 - a\}$. Further suppose that bound in (9) holds for $i = 4$. Then we have the following asymptotic expansion*

$$\mathfrak{I} = \frac{g(x_0)e(f(x_0)) + 1/8}{\sqrt{f''(x_0)}} + \left(\frac{\Omega_f^4}{\Theta_f^2 \kappa^3} + \frac{\Omega_f}{\Theta_f^{3/2}} + \frac{\Omega_f^3}{\Theta_f^{3/2}\Omega_g^2}\right).$$

Proof. See Theorem 1 and Theorem 2 of [Huxley 1994]. □

4. Proof of Theorem 1.2

In order to prove Theorem 1.2, we shall establish the following bound.

Proposition 4.1. *We have*

$$S(N) \ll \begin{cases} N^{1+\varepsilon} & \text{if } 1 \le N \ll p^{2r/3+\varepsilon}, \\ N^{1/2} p^{1/2(r - \lfloor r/3 \rfloor)+\varepsilon} & \text{if } p^{2r/3} \ll N \ll p^{r+\varepsilon}. \end{cases}$$

Application of the circle method. We first separate the oscillation of Fourier coefficients $\lambda_f(n)$ and $\chi(n)$ using delta symbol. We write

$$S(N) := \sum_{m,n=1}^{\infty}\sum \lambda_f(n)\chi(m)\delta(n-m)V\left(\frac{n}{N}\right)V_1\left(\frac{m}{N}\right),$$

where $V_1(y)$ is another smooth function, supported on the interval $\left[\frac{1}{2}, 3\right]$, $V_1(y) \equiv 1$ for $y \in [1,2]$ and satisfies $y^j V^{(j)}(y) \ll_j 1$. To analyze sum $S(N)$ we use a conductor lowering mechanism (see [Munshi 2015a] for a discrete version of the conductor lowering method and [Munshi 2015b] for the integral version). The integral equation $n = m$ is equivalent to the congruence $n \equiv m \pmod{p^\ell}$ and the integral equation $(n-m)/p^\ell = 0$, $\ell < r$. This lowers the conductor, as modulus p^ℓ is already present in the character χ. We obtain

$$S(N) := \sum_{\substack{m,n=1 \\ p^\ell|(n-m)}}^{\infty}\sum \lambda_f(n)\chi(m)\delta\left(\frac{n-m}{p^\ell}\right)V\left(\frac{n}{N}\right)V_1\left(\frac{m}{N}\right),$$

Now using Lemma 3.3 for the expression of $\delta(n)$, we have $S(N)=S^{+}(N)+S^{-}(N)$, with

$$S^{\pm}(N)$$
$$=\int_0^1 \sum_{1\le q\le Q<q\le q+Q}\sideset{}{^\star}\sum \frac{1}{aq}\sum_{\substack{m,n=1\\ p^\ell|(n-m)}}^{\infty}\lambda_f(n)\chi(m)e\left(\pm\frac{\bar{a}(n-m)/p^\ell}{q}\mp\frac{x(n-m)/p^\ell}{aq}\right)V\left(\frac{n}{N}\right)V_1\left(\frac{m}{N}\right)dx.$$

We choose $Q=(N/p^\ell)^{1/2}$. We detect the congruence relation $n\equiv m \pmod{p^\ell}$ in the above expression using an exponential sum. We have

$$S^{\pm}(N)=\int_0^1 \sum_{1\le q\le Q<q\le q+Q}\sideset{}{^\star}\sum \frac{1}{aqp^\ell}\sum_{b(p^\ell)}\sum_{m,n=1}^{\infty}$$
$$\lambda_f(n)\chi(m)e\left(\pm\frac{(\bar{a}+bq)(n-m)}{p^\ell q}\right)e\left(\mp\frac{x(n-m)}{ap^\ell q}\right)V\left(\frac{n}{N}\right)V_1\left(\frac{m}{N}\right)dx.$$

We will now analyze the sum $S^{+}(N)$ (analysis of $S^{-}(N)$ is similar). We rearrange the sum as

$$S^{+}(N)=\int_0^1 \sum_{1\le q\le Q<q\le q+Q}\sideset{}{^\star}\sum \frac{1}{aqp^\ell}\sum_{b(p^\ell)}\left\{\sum_{n=1}^{\infty}\lambda_f(n)e\left(\frac{(\bar{a}+bq)n}{p^\ell q}\right)e\left(\frac{xn}{p^\ell aq}\right)V\left(\frac{n}{N}\right)\right\}$$
$$\left\{\sum_{m=1}^{\infty}\chi(m)e\left(-\frac{(\bar{a}+bq)m}{p^\ell q}\right)e\left(\frac{-mx}{p^\ell aq}\right)V_1\left(\frac{m}{N}\right)\right\}dx. \quad (10)$$

Applying Poisson summation formula. We shall apply the Poisson summation formula to the sum over m in (10) as follows. Writing $m=\beta+cp^r q$, $c\in\mathbb{Z}$ and then applying the Poisson summation formula to sum over c, we have

$$\sum_{m=1}^{\infty}\chi(m)e\left(-\frac{(\bar{a}+bq)m}{p^\ell q}\right)e\left(\frac{-mx}{p^\ell aq}\right)V_1\left(\frac{m}{N}\right)$$
$$=\sum_{\beta(p^r q)}\chi(\beta)e\left(-\frac{(\bar{a}+bq)\beta}{p^\ell q}\right)\sum_{m\in\mathbb{Z}}\int_{\mathbb{R}}V_1\left(\frac{\beta+yp^r q}{N}\right)e\left(\frac{-(\beta+yp^r q)x}{p^\ell aq}\right)e(-my)\,dy.$$

We now substitute the change of variable $(\beta+yp^r q)/N=z$ to obtain

$$=\frac{N}{p^r q}\sum_{m\in\mathbb{Z}}\left\{\sum_{\beta(p^r q)}\chi(\beta)e\left(-\frac{(\bar{a}+bq)\beta}{p^\ell q}+\frac{m\beta}{p^r q}\right)\right\}\int_{\mathbb{R}}V_1(y)e\left(\frac{-Nxz}{p^\ell aq}\right)e\left(\frac{-Nmz}{p^r q}\right)dz$$
$$:=\frac{N}{p^r q}\mathcal{C}(b,q)\mathcal{I}(x,q,m), \quad (11)$$

where $\mathcal{C}(b,q)$ is the character sum and $\mathcal{I}(x,q,m)$ is the integral in the above expression. We now first evaluate the character sum in the following subsection.

Evaluation of the character sum. Writing $q = p^{r_1}q'$ with $(p, q') = 1$, the character sum in (11) can be written as

$$\mathcal{C}(b,q) = \sum_{\beta(p^{r+r_1}q')} \chi(\beta) e\left(\frac{-(\bar{a}+bq)\beta}{p^{\ell+r_1}q'} + \frac{m\beta}{p^{r+r_1}q'}\right).$$

Writing $\beta = \alpha_1 q'\overline{q'} + \alpha_2 p^{r+r_1}\overline{p^{r+r_1}}$, the above character sum splits as

$$\sum_{\alpha_1(p^{r+r_1})} \chi(\alpha_1) e\left(\frac{-(\bar{a}+bq)\alpha_1\overline{q'}}{p^{\ell+r_1}} + \frac{m\alpha_1\overline{q'}}{p^{r+r_1}}\right) \sum_{\alpha_2(q')} e\left(\frac{-(\bar{a}+bq)\alpha_2\overline{p^{r+r_1}}\,p^{r-\ell}}{q'} + \frac{m\alpha_2\overline{p^{r+r_1}}}{q'}\right).$$

Again, writing $\alpha_1 = \beta_1 p^r + \beta_2$, where β_2 is modulo p^r and β_1 modulo p^{r_1}, we obtain

$$\mathcal{C}(b,q) = \sum_{\beta_2(p^r)} \chi(\beta_2) e\left(\frac{-(\bar{a}+bq)p^{r-\ell}\beta_2\overline{q'}}{p^{r+r_1}} + \frac{m\beta_2\overline{q'}}{p^{r+r_1}}\right) \sum_{\beta_1(p^{r_1})} e\left(\frac{-(\bar{a}+bq)\beta_1\overline{q'}\,p^{r-\ell}}{p^{r_1}} + \frac{m\beta_1\overline{q'}}{p^{r_1}}\right)$$
$$\sum_{\alpha_2(q')} e\left(\frac{-(\bar{a}+bq)\alpha_2\overline{p^{r+r_1}}\,p^{r-\ell}}{q'} + \frac{m\alpha_2\overline{p^{r+r_1}}}{q'}\right).$$

From the last two exponential sums, we obtain the congruence relations $m - \bar{a}p^{r-\ell} \equiv 0 \pmod{p^{r_i}}$ and $m - \bar{a}p^{r-\ell} \equiv 0 \pmod{p'}$. Since we have $q = q'p^{r_1}$, we obtain the congruence relation $m - \bar{a}p^{r-\ell} \equiv 0 \pmod 1$, from which $a \pmod q$ can be determined. The sum over β_2 can be written as

$$\sum_{\beta_2(p^r)} \chi(\beta_2) e\left(\frac{(m-(\bar{a}+bq)p^{r-\ell})\beta_2\overline{q'}}{p^{r+r_1}}\right) = \chi(q')\bar{\chi}\left(\frac{m-(\bar{a}+bq)p^{r-\ell}}{p^{r_1}}\right) \sum_{\beta_2(p^r)} \chi(\beta_2) e\left(\frac{\beta_2}{p^r}\right),$$

as $p^{r_1} \mid (m-(\bar{a}+bq)p^{r-\ell})$. We record this into following lemma.

Lemma 4.2. *Let $\mathcal{C}(b,q)$ be as given in (11). We have*

$$\mathcal{C}(b,q) = \begin{cases} q\chi(q')\bar{\chi}\left(\frac{m-(\bar{a}+bq)p^{r-\ell}}{p^{r_1}}\right)\tau_\chi & \text{if } m \equiv \bar{a}p^{r-\ell} \mod q \\ 0 & \text{otherwise,} \end{cases}$$

where $q = q'p^{r_1}$ and τ_χ denotes the Gauss sum.

For simplicity of notation we assume that $q = q'(r_1 = 0)$, as the number of r_1 are bounded by $O(\log p^r)$. Next we consider the integral in (11). Integrating by parts j-times and using $V_1^{(j)}(y) \ll 1$, we have

$$\mathcal{I}(x,q,m) \ll \left(\frac{Nx}{p^\ell aq} + \frac{Nm}{p^r q}\right)^{-j}.$$

We observe that this integral is negligibly small if

$$\left|\frac{Nx}{p^\ell aq} + \frac{Nm}{p^r q}\right| \gg N^\varepsilon.$$

From the above inequality we obtain the effective range of x (in the integral originating from the circle method) as

$$\left|\frac{Nx}{p^\ell aq} + \frac{Nm}{p^r q}\right| \ll N^\varepsilon \Rightarrow \left|x + \frac{ma}{p^{r-\ell}}\right| \ll \frac{qN^\epsilon}{Q}, \tag{12}$$

as $a \asymp Q$ and $N/p^{\ell} = Q^2$. Again integrating by parts, taking $V_1(y)e((-Nxy)/(p^{\ell}aq))$ as the first function, we obtain

$$\mathcal{I}(x,q,m) \ll \left(1+\frac{Nx}{p^{\ell}aq}\right)^j \left(\frac{p^r q}{Nm}\right)^j.$$

Hence the integral is negligibly small if $m \gg (p^r Q N^{\varepsilon})/N$. After a first application of the Poisson summation formula we are left with the following expression for $S^+(N)$:

Lemma 4.3. *Let $S^+(N)$ be as given in (10). We have*

$$S^+(N) = \int_{x \ll (qN^{\epsilon})/Q} \sum_{1\le q\le Q} \frac{1}{aqp^{\ell}} \sum_{b(p^{\ell})} \left\{ \sum_{n=1}^{\infty} \lambda_f(n) e\left(\frac{(\bar{a}+bq)n}{p^{\ell}q}\right) e\left(\frac{xn}{p^{\ell}aq}\right) V\left(\frac{n}{N}\right) \right\} \left\{ \frac{\tau_{\chi}\chi(q)N}{p^r} \sum_{m \ll \frac{N^{\varepsilon}Qp^r}{N}} \bar{\chi}(m-(\bar{a}+bq)p^{r-\ell}) \mathcal{I}(x,q,m) \right\} dx + O_A(p^{-A}), \quad (13)$$

for any real $A > 0$.

Estimating trivially at this stage, we have

$$S^+(N) \ll \sum_{1\le q\le Q} \frac{1}{aqp^{\ell}} \sum_{b(p^{\ell})} \sum_{n=N}^{2N} |\lambda_f(n)| \frac{|\tau_{\chi}|N}{p^r} \sum_{m \ll \frac{N^{\varepsilon}Qp^r}{N}} 1 \ll Np^{r/2}.$$

Hence we are able to save $N/p^{r/2}$ from the first application of the Poisson summation formula.

Applying Voronoi summation formula. At this stage we need to differentiate between the holomorphic and the Maass case as the integral transforms appearing in the Voronoi summation formula are different. Nevertheless they are essentially same as far as our argument is concerned. Below we will present the details for the Maass case which is traditionally considered to be harder.

We have $(\bar{a}+bq, q) = 1$. Given a, there exists at most one b mod p^{ℓ} such that $\bar{a}+bq \equiv 0 \pmod{q^{\ell}}$. For the rest of b we apply the Voronoi summation formula to the sum over n as follows (The case where $\bar{a}+bq \equiv 0 \pmod{q^{\ell}}$ is similar and even simpler. We first write $\bar{a}+bq = p^{\ell}q_1$, and then apply the Voronoi summation formula, which gives us more saving as the conductor is now smaller than qp^{ℓ}. Also we have a savings of a whole summation over b modulo p^{ℓ}). We substitute $g(n) = e(-nx/p^{\ell}aq)V(n/N)$ in Lemma 3.1 to get

$$\sum_{n=1}^{\infty} \lambda_f(n) e\left(\frac{(\bar{a}+bq)n}{p^{\ell}q}\right) e\left(\frac{xn}{p^{\ell}aq}\right) V\left(\frac{n}{N}\right) = \frac{1}{p^{\ell}q} \sum_{\pm} \sum_{n\ge 1} \lambda_f(\mp n) e\left(\pm n \frac{\overline{\bar{a}+bq}}{p^{\ell}q}\right) H^{\pm}\left(\frac{n}{q^2}, \frac{Nx}{aq}\right),$$

where

$$H^{-}\left(\frac{n}{q^2}, \frac{Nx}{p^{\ell}aq}\right) = \int_0^{\infty} e\left(-\frac{xy}{p^{\ell}aq}\right) V\left(\frac{y}{N}\right) \{Y_{2i\nu} + Y_{-2i\nu}\} \left(\frac{4\pi\sqrt{ny}}{p^{\ell}q}\right) dy$$

(we have similar expression for $H^+(x, y)$). Substituting $y/N = z$, we have

$$H^-\left(\frac{n}{q^2}, \frac{Nx}{P_1aq}\right) = N\int_0^\infty e\left(-\frac{Nxz}{p^\ell aq}\right)V\left(z\right)\{Y_{2i\nu}+Y_{-2i\nu}\}\left(\frac{4\pi\sqrt{nNz}}{p^\ell q}\right)dz := N\mathcal{J}(x,n,q), \quad (14)$$

where $\mathcal{J}(x, n, q)$ denotes the integral in above equation. Pulling out the oscillations, we have the following asymptotic formulae for Bessel functions (see [Kowalski et al. 2002, Lemma C.2]):

$$Y_{\pm 2i\nu}(x) = e^{ix}U_{\pm 2i\nu}(x) + e^{-ix}\overline{U}_{\pm 2i\nu}(x) \quad \text{and} \quad |x^k K_\nu^{(k)}(x)| \ll_{k,\nu} \frac{e^{-x}(1+\log|x|)}{(1+x)^{1/2}}, \quad (15)$$

where the function $U_{\pm 2i\nu}(x)$ satisfies,

$$x^j U_{\pm 2i\nu}^{(j)}(x) \ll_{j,\nu,k} (1+x)^{-1/2}. \quad (16)$$

We also have $J_k(x) = e^{ix}W_k(x) + e^{-ix}\overline{W}(x)$, where

$$x^j W_k^{(j)}(x) \ll_j \frac{1}{(1+x)^{1/2}}.$$

Substituting the above decomposition for $Y_{\pm 2i\nu}(x)$, the first term of the integral in (14) is given by (estimation of the second term is similar)

$$\mathcal{J}^\pm(x,n,q) := \int_0^\infty e^{i(-(2\pi Nxy)/(p^\ell aq)\pm i(4\pi\sqrt{nNy})/(p^\ell q))}V(y)U_{2i\nu}^\pm\left(\frac{4\pi\sqrt{nNy}}{p^\ell q}\right)dy, \quad (17)$$

where we have denoted $U^+(y) := U(y)$ and $U^-(y) = \overline{U}(y)$. The integral $\mathcal{J}^-(x, n, q)$ has no stationary point. By (8), $\mathcal{J}^-(x, n, q)$ is negligibly small. For $\mathcal{J}^+(x, n, q)$ we apply the second statement of Lemma 3.4 with

$$f(y) = -\frac{2\pi Nxy}{p^\ell aq} + i\frac{4\pi\sqrt{nNy}}{p^\ell q} \quad \text{and} \quad g(y) = V(y)U_{2i\nu}\left(\frac{4\pi\sqrt{nNy}}{p^\ell q}\right).$$

We have

$$f'(y) = -\frac{2\pi Nx}{p^\ell aq} + \frac{2\pi\sqrt{nN}}{\sqrt{y}p^\ell q}, \quad f''(y) = -\frac{\pi\sqrt{nN}}{y^{3/2}p^\ell q}.$$

We observe that

$$|F''(y_0)| \asymp \frac{\sqrt{nN}}{p^\ell q},$$

where $y_0 = na^2/(Nx^2)$ is the stationary point, which is $y_0 \asymp 1$ as $V(y)$ is supported on the interval $[1, 2]$. Using $U_{\pm 2i\nu}(x) \ll_\nu (1+x)^{-1/2}$, and applying the second statement of Lemma 3.4, we obtain

$$\mathcal{J}(x,n,q) \ll \frac{p^\ell q}{\sqrt{nN}}, \quad (18)$$

where $\mathcal{J}(x, n, q)$ is given in (14). Also, integrating by parts we have

$$\mathcal{J}(x,n,q) \ll_j \left(\frac{Nx}{p^\ell aq}+1\right)^j\left(\frac{p^\ell q}{\sqrt{nN}}\right)^j.$$

The integral is negligibly small if (note that $x \ll qN^{\epsilon}/Q$)

$$\frac{p^{\ell}q}{\sqrt{nN}} \ll p^{-\varepsilon} \Rightarrow n \gg p^{\varepsilon}p^{\ell}.$$

We record this result in the following lemma. After applying the Poisson and the Voronoi summation formula we have the following expression for $S^{+}(N)$.

Lemma 4.4. *We have*

$$S^{+}(N) = \int_{x \ll (qN^{\varepsilon})/Q} \sum_{1\le q\le Q} \frac{1}{qp^{\ell}} \sideset{}{^\star}\sum_{b(p^{\ell})} \frac{\tau_{\chi}\chi(q)N}{p^{r}} \sum_{m \ll \frac{Qp^{r}p^{\varepsilon}}{N}} \tfrac{1}{a}\overline{\chi}(m-(\bar{a}+bq)p^{r-\ell})\mathcal{I}(x,q,m)$$
$$\left\{\frac{N}{p^{\ell}q}\sum_{\pm}\sum_{n\ll p^{\ell}p^{\varepsilon}} \lambda_f(\mp n)e\left(\pm n\frac{\overline{\bar{a}+bq}}{p^{\ell}q}\right)\mathcal{J}(x,n,q)\right\}dx + O_A(p^{-A}).$$

Estimating trivially, we have (assuming square-root cancellation in the character sum over b and Lemma 3.2):

$$S^{+}(N) \ll \sum_{1\le q\le Q}\frac{1}{qp^{\ell}}\frac{|\tau_{\chi}|N}{p^{r}}\sum_{m\ll\frac{p^{\varepsilon}Qp^{r}}{N}}\frac{|\mathcal{I}(x,q,m)|}{a}\frac{N}{p^{\ell}q}\sum_{n\ll p^{\ell}N^{\varepsilon}}|\lambda_f(\mp n)\mathcal{J}(x,n,q)|$$
$$\left|\sideset{}{^\star}\sum_{b(p^{\ell})} e\left(\pm n\frac{\overline{\bar{a}+bq}}{p^{\ell}q}\right)\overline{\chi}\left(m-(\bar{a}+bq)p^{r-\ell}\right)\right|$$
$$\ll \frac{1}{ap^{\ell}}\frac{p^{r/2}N}{p^{r}}\frac{N^{\varepsilon}Qp^{r}}{N}\times\frac{N}{p^{\ell}q}\sum_{n\ll p^{\ell}}\frac{p^{\ell}q}{\sqrt{nN}}\times p^{\ell/2}$$
$$\ll N^{\varepsilon}\sqrt{N}p^{r/2}.$$

This shows that we are on the boundary. To obtain an additional saving, we shall now apply the Cauchy–Schwarz inequality to the summation over n and then apply the Poisson summation formula. Interchanging the order of summation, we have

$$S^{+}(N) = \frac{N^{2}\tau_{\chi}}{p^{r+2\ell}}\sum_{n\ll p^{\ell}p^{\varepsilon}}\lambda_f(n)\hat{S}_1(n) + O_A(p^{-A}), \tag{19}$$

where

$$\hat{S}_1(n) = \int_{x\ll(qN^{\epsilon})/Q}\sum_{1\le q\le Q}\sideset{}{^\star}\sum_{b(p^{\ell})}\sum_{m\ll\frac{p^{\varepsilon}Qp^{r}}{N}}$$
$$\frac{\chi(q)}{aq^{2}}\overline{\chi}(m-(\bar{a}+bq)p^{r-\ell})e\left(-n\frac{\overline{a+bq}\bar{q}}{p^{\ell}}\right)e\left(-n\frac{p^{r}\overline{p^{2\ell}}\overline{m}}{q}\right)\mathcal{I}(x,q,m)\mathcal{J}(x,n,q)\,dx.$$

5. Proof of Theorem 1.2 — Conclusion

In the previous section we have completed the first three steps of the proof as given in the short sketch. As expected we are at the threshold and any saving will yield subconvexity. We now apply Cauchy–Schwarz to escape from the "trap of involution" and to get rid of the Fourier coefficients.

Applying Cauchy inequality. We split the summation over n into dyadic sum. Applying the Cauchy–Schwarz inequality on the summation over n in (19) and using Lemma 3.2, we have

$$S^+(N) \ll \frac{N^2|\tau_\chi|}{p^{r+2\ell}} \sum_{\substack{L\ll P_1\\ L\text{-dyadic}}} \Big\{\sum_{n\ll L}|\lambda_f(n)|^2\Big\}^{1/2} \Big\{\sum_{n\in\mathbb{Z}}|\hat{S}_1(n)|^2 U\Big(\frac{n}{L}\Big)\Big\}^{1/2}$$
$$\ll \frac{N^2|\tau_\chi|}{p^{r+2\ell}} \sum_{\substack{L\ll P_1\\ L\text{-dyadic}}} L^{1/2}\{\hat{S}_2(L)\}^{1/2}, \tag{20}$$

where $P_1 = p^{\ell+\varepsilon}$ and $\hat{S}_2(L)$ is given by (opening the absolute square and pushing the summation over n inside):

$$\hat{S}_2(L) := \int_{x\ll(qN^\epsilon)/Q}\int_{x'\ll(q'N^\epsilon)/Q} \sum_{m\ll\frac{N^\varepsilon Qp^r}{N}} \sum_{m'\ll\frac{N^\varepsilon Qp^r}{N}} \sum_{1\le q\le Q}\sum_{1\le q'\le Q} \frac{\chi(q)}{aq^2}\frac{\overline{\chi}(q')}{a'q'^2} \sideset{}{^\star}\sum_{b(p^\ell)} \sideset{}{^\star}\sum_{b'(p^\ell)}$$
$$\overline{\chi}(m-(\bar{a}+bq)p^{r-\ell})\chi(m'-(\bar{a}+b'q')p^{r-\ell})\mathfrak{I}(x,q,m)\mathfrak{I}(x,q',m')\mathcal{T}\,dx\,dx', \tag{21}$$

where

$$\mathcal{T} := \sum_{n\in\mathbb{Z}} e\bigg(n\bigg\{-\frac{p^r\,\overline{p^{2\ell}}\,\overline{m}}{q} + \frac{p^r\,\overline{p^{2\ell}}\,\overline{m'}}{q'} - \frac{\overline{a+bq}\,\bar{q}}{p^\ell} + \frac{\overline{a+b'q'}\,\overline{q'}}{p^\ell}\bigg\}\bigg) U\Big(\frac{n}{L}\Big)\mathcal{J}(x,n,q)\mathcal{J}(x,n,q'). \tag{22}$$

Second application of Poisson summation formula. We write a smooth bump function

$$U(n/L)\mathcal{J}(x,n,q)\mathcal{J}(x,n,q') := U_1(n/L),$$

where $\mathcal{J}(x,n,q)$ is as given in (14). Writing $n = \alpha + qq'p^\ell c$, $c\in\mathbb{Z}$ and applying Poisson summation formula to sum over c, we have

$$\mathcal{T} := \sum_{\alpha(qq'p^\ell)} e\bigg(\alpha\bigg\{-\frac{p^r\,\overline{p^{2\ell}}\,\overline{m}}{q} + \frac{p^r\,\overline{p^{2\ell}}\,\overline{m'}}{q'} - \frac{\overline{a+bq}\,\bar{q}}{p^\ell} + \frac{\overline{a+bq'}\,\overline{q'}}{p^\ell}\bigg\}\bigg) \int_{\mathbb{R}} \sum_{n\in\mathbb{Z}} U_1\Big(\frac{\alpha+yqq'p^\ell}{L}\Big) e(-ny)\,dy.$$

We now apply the change of variable $(\alpha + yqq'p^\ell)/L = z$ to get

$$\mathcal{T} = \frac{L}{qq'p^\ell}\sum_{n\in\mathbb{Z}}\sum_{\alpha(qq'p^\ell)} e\bigg(\alpha\bigg\{-\frac{p^r\,\overline{p^{2\ell}}\,\overline{m}}{q} + \frac{p^r\,\overline{p^{2\ell}}\,\overline{m'}}{q'} - \frac{\overline{a+bq}\,\bar{q}}{p^\ell} + \frac{\overline{a+bq'}\,\overline{q'}}{p^\ell} + \frac{n}{qq'p^\ell}\bigg\}\bigg)$$
$$\int_{\mathbb{R}} U_1\Big(y\Big) e\Big(-\frac{nLy}{qq'p^\ell}\Big)\,dy. \tag{23}$$

We have $U_1(y) = U(y)\mathcal{J}(x, Ly, q)\mathcal{J}(x, Ly, q')$. From the expression of $\mathcal{J}(x, Lu, q)$ in (14) (note that after change of variable we have $u \asymp 1$) and (16), we have

$$\begin{aligned}\frac{\partial}{\partial u}\mathcal{J}(x, Lu, q) &= \int_0^\infty e\left(-\frac{Nxy}{p^\ell aq}\right)V(y)\frac{\partial}{\partial u}\{Y_{2i\nu} + Y_{-2i\nu}\}\left(\frac{4\pi\sqrt{LuNy}}{p^\ell q}\right)dy\\ &= \int_0^\infty e\left(-\frac{Nxy}{p^\ell aq}\right)V(y)\frac{1}{u}\frac{4\pi\sqrt{LuNy}}{p^\ell q}\{Y'_{2i\nu} + Y'_{-2i\nu}\}\left(\frac{4\pi\sqrt{LuNy}}{p^\ell q}\right)dy\\ &\ll 1.\end{aligned}$$

This shows that there is no oscillation in the function $\mathcal{J}(x, Ln, q)$ with respect to variable n. Similarly, higher order derivatives of $\mathcal{J}(x, Ln, q)$ with respect to u are bounded. Also from (18) we have

$$\begin{aligned}\int_{\mathbb{R}} U_1(y)e\left(-\frac{nLy}{qq'p^\ell}\right)dy &= \int_{\mathbb{R}} U(y)\mathcal{J}(x, Ly, q)\mathcal{J}(x, Ly, q')e\left(-\frac{nLy}{qq'p^\ell}\right)dy\\ &\ll \frac{p^\ell q}{\sqrt{LN}}\frac{p^\ell q'}{\sqrt{LN}}\int_1^2 U(y)\,dy\\ &\ll \frac{p^{2\ell}qq'}{LN},\end{aligned} \tag{24}$$

as $U(y)$ is supported on the interval $[1, 2]$.

Integrating by parts taking $U_1(y)$ as first function, we observe that the integral in (23) is negligible if $n \gg p^\varepsilon qq'p^\ell/L$. Evaluating the above character sum we get the following congruence relation:

$$-p^r\,\overline{p^{2\ell}}\,\overline{m}\,p^\ell q' + p^r\,\overline{p^{2\ell}}\,\overline{m'}\,p^\ell q - \overline{a+bq}\,\bar{q}qq' + \overline{a+b'q'}\,\overline{q'}qq' + n \equiv 0 \pmod{qq'p^\ell}.$$

Here $\bar{q}$ and $\overline{q'}$ are the inverses of q and q' modulo p^ℓ respectively. We solve the above congruence modulo p^ℓ and modulo qq' respectively to obtain

$$-\overline{a+bq}\,q' + \overline{a+b'q'}\,q + n \equiv 0 \pmod{p^\ell} \quad\text{and}\quad -p^r\,\overline{p^{2\ell}}\,\overline{m}\,p^\ell q' + p^r\,\overline{p^{2\ell}}\,\overline{m'}\,p^\ell q + n \equiv 0 \pmod{qq'}. \tag{25}$$

Writing $n = -p^r\,\overline{p^{2\ell}}\,\overline{m}\,p^\ell q' + p^r\,\overline{p^{2\ell}}\,\overline{m'}\,p^\ell q + jqq'$, we observe that the number of n satisfying the above congruence relation is same as the number of j's. Since we also have $n \ll N^\varepsilon qq'p^\ell/L$, we conclude $j \ll N^\varepsilon$. Hence the number of solutions of n satisfying the above congruence relation modulo qq', and $n \ll qq'N^\varepsilon p^\ell/L$ is bounded by $N^\varepsilon p^\ell/L$. For congruence relation modulo p^ℓ in the above equation, we substitute the change of variable $a + bq = \alpha$ and $a + b'q' = \alpha'$ to obtain

$$\bar{\alpha}q' + \overline{\alpha'}q + n \equiv 0 \pmod{p^\ell}. \tag{26}$$

We record the bound for $\mathcal{T}$ in the following lemma:

Lemma 5.1. *Let $\mathcal{T}$ be as given in (22). We have*

$$\mathcal{T} = L\sum_{n \ll p^\varepsilon qq'p^\ell/L}^{\dagger}\int_{\mathbb{R}} U_1(y)e\left(-\frac{nLy}{qq'p^\ell}\right)dy,$$

where † in the above summation denotes that n satisfies the congruence relation given in (25).

Substituting the bound for $\mathcal{T}$ in (21) we obtain

$$\hat{S}_2(L) = L\int_{x\ll(qN^\epsilon)/Q}\int_{x'\ll(q'N^\epsilon)/Q}\sum_{m\ll\frac{N^\varepsilon Qp^r}{N}}\sum_{m'\ll\frac{N^\varepsilon Qp^r}{N}}\sum_{1\le q\le Q}\sum_{1\le q'\le Q}\frac{\chi(q)}{aq^2}\frac{\overline{\chi}(q')}{a'q'^2}\sideset{}{^\dagger}\sum_{n\ll p^\varepsilon qq'p^\ell/L}\sideset{}{^\star}\sum_{\alpha(p^\ell)}\sideset{}{^\star}\sum_{\alpha'(p^\ell)}$$

$$\overline{\chi}(m-\alpha p^{r-\ell})\chi(m'-\alpha' p^{r-\ell})\mathfrak{I}(x,q,m)\mathfrak{I}(x,q',m')\int_{\mathbb{R}}U_1(y)e\left(-\frac{nLy}{qq'p^\ell}\right)dy\,dx\,dx'$$

$$:= \hat{S}_2(D)+\hat{S}_2(ND), \tag{27}$$

where α and α' are related by the congruence relation given in (26), and $\hat{S}_2(D)$ (respectively $\hat{S}_2(ND)$) is contribution of the diagonal terms (respectively the off-diagonal terms). The contribution of the diagonal terms ($\alpha=\alpha'$, $m=m'$ and $q=q'$) is bounded by (using $\mathfrak{I}(x,q,m)\ll 1$, bound from the (24) and sum over n satisfying the congruence relation given in (25) is bounded by $p^\epsilon p^\ell/N$):

$$\hat{S}_2(D) \ll L\int_{x\ll(qN^\epsilon)/Q}\int_{x'\ll(q'N^\epsilon)/Q}\sum_{m\ll\frac{Qp^r}{N}}\sum_{1\le q\le Q}\frac{1}{q^4}\sideset{}{^\dagger}\sum_{n\ll p^\varepsilon qq'p^\ell/L}\sum_{\alpha(p^\ell)}\frac{|\mathfrak{I}(x,q,m)|^2}{a^2}\int_{\mathbb{R}}|U_1(y)|\,dy\,dx\,dx'$$

$$\ll LN^\varepsilon\int_{x\ll q/Q}\frac{Qp^r}{N}\sum_{1\le q\le Q}\sideset{}{^\dagger}\sum_{n\ll p^\varepsilon qq'p^\ell/N}\frac{1}{a^2q^4}p^\ell\frac{p^{2\ell}q^2}{LN}\,dx$$

$$\ll \frac{p^{3\ell}N^\varepsilon}{N}\frac{Qp^r}{N}\frac{p^\ell}{L}\sum_{1\le q\le Q}\frac{1}{a^2q^2}\times\frac{q}{Q}$$

$$\ll \frac{p^{3\ell}N^\varepsilon}{Q^2N}\frac{p^r}{N}\frac{p^\ell}{L}, \tag{28}$$

as $a\asymp Q$. Substituting the value of α' from the congruence relation given in (26), we see that the contribution of the off-diagonal term is given by:

$$\hat{S}_2(ND) = L\int_{x\ll(qN^\epsilon)/Q}\int_{x'\ll(q'N^\epsilon)/Q}\sum_{m\ll Q}\sum_{m'\ll Q}\sum_{1\le q\le Q}\sum_{1\le q'\le Q}\frac{\chi(q)}{aq^2}\frac{\overline{\chi}(q')}{a'q'^2}\sideset{}{^\dagger}\sum_{n\ll p^\varepsilon qq'p^\ell/N}$$

$$\left\{\sideset{}{^\star}\sum_{\alpha(p^\ell)}\overline{\chi}(m-\alpha p^{r-\ell})\chi(m'+\alpha q\overline{n+q'}p^{r-\ell})\right\}\mathfrak{I}(x,q,m)\mathfrak{I}(x,q',m')$$

$$\int_{\mathbb{R}}U_1(y)e\left(-\frac{nLy}{qq'p^\ell}\right)dy\,dx\,dx'. \tag{29}$$

Next we evaluate the exponential sum in the above equation.

Evaluation of the character sum. In this subsection we shall prove the following lemma

Lemma 5.2. *Let $\mathcal{A}$ be the character sum given by*

$$\mathcal{A} := \sideset{}{^\star}\sum_{\alpha(p^\ell)}\overline{\chi}(m-\alpha p^{r-\ell})\chi(m'+\alpha q\overline{n+q'}p^{r-\ell}),$$

with $\ell = 2\lfloor\frac{r}{3}\rfloor$. We have

$$\mathcal{A} \ll p^{\ell/2+\epsilon}.$$

Proof. Applying the change of variable $\alpha = \alpha_1 p^{\ell/2} + \alpha_2$ where α_1 and α_2 run over residue classes modulo $p^{\ell/2}$, the above character sum reduces to

$$\begin{aligned}\mathcal{A} &= \sideset{}{^\star}\sum_{\alpha_2(p^{\ell/2})}\sideset{}{^\star}\sum_{\alpha_1(p^{\ell/2})} \bar{\chi}(m-\alpha_2 p^{r-\ell}-\alpha_1 p^{2(r-\ell)})\chi(m'+(\alpha_1 p^{r-\ell}+\alpha_2)q\overline{n+q'}p^{r-\ell})\\ &= \sideset{}{^\star}\sum_{\alpha_2(p^{\ell/2})}\sideset{}{^\star}\sum_{\alpha_1(p^{\ell/2})} \chi\{(m'+q\overline{n+q'}p^{r-\ell}\alpha_2+q\overline{n+q'}\alpha_1 p^{2(r-\ell)})(\overline{m-\alpha_2 p^{r-\ell}}+(\overline{m-\alpha_2 p^{r-\ell}})^2\alpha_1 p^{2(r-\ell)})\},\end{aligned}$$

as $\overline{m-\alpha_2 p^{r-\ell}-\alpha_1 p^{2(r-\ell)}} = \overline{m-\alpha_2 p^{r-\ell}} + (\overline{m-\alpha_2 p^{r-\ell}})^2\alpha_1 p^{2(r-\ell)} \pmod{p^r}$. Which reduces to

$$\begin{aligned}\mathcal{A} &= \sideset{}{^\star}\sum_{\alpha_2(p^{\ell/2})}\sideset{}{^\star}\sum_{\alpha_1(p^{\ell/2})} \chi(A(\alpha_2)+B(\alpha_2)\alpha_1 p^{2(r-\ell)})\\ &= \sideset{}{^\star}\sum_{\alpha_2(p^{\ell/2})}\chi(A(\alpha_2))\sideset{}{^\star}\sum_{\alpha_1(p^{\ell/2})} \chi(1+\overline{A(\alpha_2)}B(\alpha_2)\alpha_1 p^{2(r-\ell)}),\end{aligned}$$

where

$$A(\alpha_2) = m'\overline{m-\alpha_2 p^{r-\ell}} + q\overline{n+q'}\alpha_2 p^{r-\ell}\overline{m-\alpha_2 p^{r-\ell}} \quad \text{and} \quad B(\alpha_2) = m'(\overline{m-\alpha_2 p^{r-\ell}})^2 + q\overline{n+q'}\,\overline{m-\alpha_2 p^{r-\ell}}.$$

Note that $(A(\alpha_2), p) = 1$, otherwise $\chi(A(\alpha_2)+B(\alpha_2)\alpha_1 p^{2r/3}) = 0$. For a fixed α_2,

$$\chi(1+\overline{A(\alpha_2)}B(\alpha_2)\alpha_1 p^{2(r-\ell)}) := \chi(1+C(\alpha_2)\alpha_1 p^{2(r-\ell)})$$

is an additive character of modulus $p^{\ell/2}$, as we have

$$\chi(1+C(\alpha_2)\alpha_1 p^{2(r-\ell)})\chi(1+C(\alpha_2)\alpha_1' p^{2(r-\ell)}) = \chi(1+C(\alpha_2)(\alpha_1+\alpha_1')p^{2(r-\ell)}),$$

as we have $4(r-\ell) \geq r$. Hence there exists an integer b (uniquely determined modulo $p^{\ell/2}$) such that

$$\chi(1+C(\alpha_2)\alpha_1 p^{2(r-\ell)}) = e\left(\frac{\alpha_1 bC(\alpha_2)}{p^\ell}\right).$$

Executing the sum over α_1 given in Lemma 5.2 we have

$$\mathcal{A} = p^{\ell/2} \sideset{}{^\star}\sum_{\substack{\alpha_2(p^{\ell/2})\\ bC(\alpha_2)\equiv 0 \ (\mathrm{mod}\ p^{\ell/2})}} \chi(A(\alpha_2)) \ll p^{\ell/2+\varepsilon}. \tag{30}$$

This concludes the proof. □

Substituting the bound for the character sum in (29) and using the bounds of $U_1(y)$ given in (24), we have

$$\begin{aligned}\hat{S}_2(ND) &\ll p^\varepsilon L \sum_{m\ll \frac{Qp^r}{N}}\sum_{m'\ll \frac{Qp^r}{N}}\sum_{1\leq q\leq Q}\sum_{1\leq q'\leq Q}\sideset{}{^\dagger}\sum_{n\ll p^\varepsilon qq'p^\ell/L} \frac{1}{aq^2}\frac{1}{a'q'^2}p^{\ell/2}\frac{p^{2\ell}qq'}{LN}\\ &\ll p^\varepsilon \frac{p^{5\ell/2}}{Q^2N}\left(\frac{Qp^r}{N}\right)^2\frac{p^\ell}{L} \ll p^\varepsilon\frac{p^{3\ell/2}}{Q^2}\left(\frac{p^r}{N}\right)^2\frac{p^\ell}{L},\end{aligned} \tag{31}$$

as $a, a' \asymp Q$, $Q^2 = N/p^\ell$ and dagger on summation over n shows that n satisfies the congruence relation modulo qq' as given in (25). Substituting the bounds for $\hat{S}_2(D)$ and $\hat{S}_2(ND)$ in (27) we have

$$\hat{S}_2(L) \ll p^\varepsilon \frac{p^\ell}{L}\left(\frac{p^{3\ell}}{Q^2N}\frac{p^r}{N} + \frac{p^{3\ell/2}}{Q^2}\frac{p^{2r}}{N^2}\right) \ll p^\varepsilon \frac{p^\ell}{LQ^2N^2} p^{\frac{3\ell}{2}} p^r (p^{3\ell/2} + p^r).$$

Substituting the bound for $\hat{S}_2(L)$ in (20) we obtain

$$\begin{aligned} S_2^+(N) &\ll p^\varepsilon \frac{N^2|\tau_\chi|}{p^{r+2\ell}} \sum_{\substack{L \ll P_1 \\ L\text{-dyadic}}} L^{1/2} \times \frac{p^{\ell/2}}{\sqrt{L}QN} p^{3\ell/4} p^{r/2} (p^{3\ell/4} + p^{r/2}) \\ &\ll p^\varepsilon \frac{N}{Qp^{3\ell/4}} (p^{3\ell/4} + p^{r/2}) \\ &\ll p^\varepsilon N^{1/2}\left(p^{\ell/2} + \frac{p^{r/2}}{p^{\ell/4}}\right) \\ &\ll p^\varepsilon N^{1/2} p^{1/2(r-\lfloor r/3\rfloor)+\varepsilon}, \end{aligned} \tag{32}$$

as $\ell = 2\left\lfloor \frac{r}{3} \right\rfloor$ and $Q = N^{1/2}/p^{\ell/2}$. This proves Proposition 4.1. □

Acknowledgement

The second author would also like to thank the Stat-Math unit, Indian Statistical Institute, Kolkata for wonderful academic atmosphere. During the work, he was supported by the Department of Atomic Energy, Government of India, NBHM post doctoral fellowship no: 2/40(15)/2016/R&D-II/5765.

References

[Aggarwal and Singh 2017] K. Aggarwal and S. K. Singh, "t-aspect subconvexity bound for $GL(2)$ L-functions", preprint, 2017. arXiv

[Blomer and Harcos 2008] V. Blomer and G. Harcos, "Hybrid bounds for twisted L-functions", *J. Reine Angew. Math.* **621** (2008), 53–79. MR Zbl

[Blomer and Milićević 2015] V. Blomer and D. Milićević, "p-adic analytic twists and strong subconvexity", *Ann. Sci. Éc. Norm. Supér.* (4) **48**:3 (2015), 561–605. MR Zbl

[Bourgain 2017] J. Bourgain, "Decoupling, exponential sums and the Riemann zeta function", *J. Amer. Math. Soc.* **30**:1 (2017), 205–224. MR Zbl

[Burgess 1963] D. A. Burgess, "On character sums and L-series, II", *Proc. London Math. Soc.* (3) **13** (1963), 524–536. MR Zbl

[Bykovskii 1996] V. A. Bykovskii, "A trace formula for the scalar product of Hecke series and its applications", *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov.* **226** (1996), 14–36. In Russian; translated in *J. Math. Sci.***89**:1 (1998), 915–932. MR Zbl

[Conrey and Iwaniec 2000] J. B. Conrey and H. Iwaniec, "The cubic moment of central values of automorphic L-functions", *Ann. of Math.* (2) **151**:3 (2000), 1175–1216. MR Zbl

[Duke et al. 1993] W. Duke, J. Friedlander, and H. Iwaniec, "Bounds for automorphic L-functions", *Invent. Math.* **112**:1 (1993), 1–8. MR Zbl

[Good 1982] A. Good, "The square mean of Dirichlet series associated with cusp forms", *Mathematika* **29**:2 (1982), 278–295. MR Zbl

[Heath-Brown 1978] D. R. Heath-Brown, "Hybrid bounds for Dirichlet L-functions", *Invent. Math.* **47**:2 (1978), 149–170. MR Zbl

[Huxley 1994] M. N. Huxley, "On stationary phase integrals", *Glasgow Math. J.* **36**:3 (1994), 355–362. MR Zbl

[Iwaniec 1997] H. Iwaniec, *Topics in classical automorphic forms*, Graduate Studies in Mathematics **17**, American Mathematical Society, Providence, RI, 1997. MR Zbl

[Iwaniec and Kowalski 2004] H. Iwaniec and E. Kowalski, *Analytic number theory*, American Mathematical Society Colloquium Publications **53**, American Mathematical Society, Providence, RI, 2004. MR Zbl

[Iwaniec and Sarnak 1999] H. Iwaniec and P. Sarnak, "Dirichlet L-functions at the central point", pp. 941–952 in *Number theory in progress* (Zakopane-Kościelisko, 1997), vol. 2, edited by K. Győry et al., de Gruyter, Berlin, 1999. MR Zbl

[Jutila 1987] M. Jutila, *Lectures on a method in the theory of exponential sums*, Tata Institute of Fundamental Research Lectures on Mathematics and Physics **80**, Springer, 1987. MR Zbl

[Kowalski et al. 2002] E. Kowalski, P. Michel, and J. VanderKam, "Rankin–Selberg L-functions in the level aspect", *Duke Math. J.* **114**:1 (2002), 123–191. MR Zbl

[Landau 1924] E. Landau, "Über die ξ-Funktion und die L-Funktionen", *Math. Z.* **20**:1 (1924), 105–125. MR JFM

[Meurman 1988] T. Meurman, "On exponential sums involving the Fourier coefficients of Maass wave forms", *J. Reine Angew. Math.* **384** (1988), 192–207. MR Zbl

[Meurman 1990] T. Meurman, "On the order of the Maass L-function on the critical line", pp. 325–354 in *Number theory* (Budapest, 1987), vol. 1, edited by K. Győry and G. Halász, Colloq. Math. Soc. János Bolyai **51**, North-Holland, Amsterdam, 1990. MR Zbl

[Milićević 2016] D. Milićević, "Sub-Weyl subconvexity for Dirichlet L-functions to prime power moduli", *Compos. Math.* **152**:4 (2016), 825–875. MR

[Munshi 2014] R. Munshi, "The circle method and bounds for L-functions, I", *Math. Ann.* **358**:1-2 (2014), 389–401. MR Zbl

[Munshi 2015a] R. Munshi, "The circle method and bounds for L-functions, II: Subconvexity for twists of GL(3) L-functions", *Amer. J. Math.* **137**:3 (2015), 791–812. MR Zbl

[Munshi 2015b] R. Munshi, "The circle method and bounds for L-functions, III: t-aspect subconvexity for GL(3) L-functions", *J. Amer. Math. Soc.* **28**:4 (2015), 913–938. MR Zbl

[Munshi 2015c] R. Munshi, "The circle method and bounds for L-functions, IV: Subconvexity for twists of GL(3) L-functions", *Ann. of Math.* (2) **182**:2 (2015), 617–672. MR Zbl

[Munshi 2018] R. Munshi, "A note on Burgess bound", pp. 273–289 in *Geometry, algebra, number theory, and their information technology applications*, edited by A. Akbary and S. Gun, Springer Proc. Math. Stat. **251**, Springer, 2018. MR

[Titchmarsh 1986] E. C. Titchmarsh, *The theory of the Riemann zeta-function*, 2nd ed., The Clarendon Press, New York, 1986. MR Zbl

[Weyl 1916] H. Weyl, "Über die Gleichverteilung von Zahlen mod. Eins", *Math. Ann.* **77**:3 (1916), 313–352. MR Zbl

Communicated by Philippe Michel
Received 2018-07-09 Revised 2019-03-12 Accepted 2019-04-10

ritabratamunshi@gmail.com *Stat-Math Unit, Indian Statistical Institute, AN Kolmogorov Bhavan, Kolkata, India*

skumar.bhu12@gmail.com *Stat-Math Unit, Indian Statistical Institute, AN Kolmogorov Bhavan, Kolkata, India*

ALGEBRA AND NUMBER THEORY 13:6 (2019)
dx.doi.org/10.2140/ant.2019.13.1415

Examples of hypergeometric twistor $\mathcal{D}$-modules

Alberto Castaño Domínguez, Thomas Reichelt and Christian Sevenheck

We show that certain one-dimensional hypergeometric differential systems underlie objects of the category of irregular mixed Hodge modules, which was recently introduced by Sabbah, and compute the irregular Hodge filtration for them. We also provide a comparison theorem between two different types of Fourier–Laplace transformation for algebraic integrable twistor $\mathcal{D}$-modules.

1. Introduction

In a series of papers Sabbah and Yu (partly joint with Esnault) [Yu 2014; Sabbah and Yu 2015; Esnault et al. 2017; Sabbah 2018] considered a so-called irregular Hodge filtration on certain cohomology groups and on certain irregular $\mathcal{D}$-modules. It can be seen as a generalization of the Hodge filtration on a mixed Hodge module in the sense of M. Saito. Geometrically, such a filtration arises by considering a version of the twisted de Rham cohomology of certain proper maps, and it plays (conjecturally) a role in Hodge theoretic mirror symmetry (see [Katzarkov et al. 2017]). Sabbah [2018] has defined a category of irregular mixed Hodge modules, which is (up to a technical equivalence) a certain subcategory of T. Mochizuki's category of (integrable) mixed twistor $\mathcal{D}$-modules. He has proved that a rigid irreducible $\mathcal{D}$-module on the projective line can be uniquely upgraded to an irregular Hodge module if and only if its formal local monodromies are unitary. Consequently, these objects come equipped with an irregular Hodge filtration and one can define irregular Hodge numbers for them. They should be seen as interesting numerical invariants attached to these differential systems, contrary to the case of arbitrary mixed twistor $\mathcal{D}$-modules, where there is no obvious way to define such numbers. In [Castaño Domínguez and Sevenheck 2019], the first and the third named author have computed that filtration and its corresponding numbers for the purely irregular hypergeometric modules, that is for systems of the form $\mathcal{D}_{\mathbb{G}_m}/\mathcal{D}_{\mathbb{G}_m}P$, where P is the operator

$$P = \prod_{i=1}^{n}(t\partial_t - \alpha_i) - t$$

The authors are partially supported by the project SISYPH: ANR-13-IS01-0001-01/02 and DFG grant SE 1114/5-1. Castaño Domínguez is also partially supported by MTM2014-59456-P, the ERDF and GI-2136. Reichelt is supported by a DFG Emmy Noether Fellowship (RE 3567/1-1).
MSC2010: primary 14F10; secondary 32C38.
Keywords: D-modules, irregular Hodge filtration, twistor D-modules, Fourier–Laplace transformation, hypergeometric D-modules.

for real numbers $\alpha_1, \dots, \alpha_n$. Let us consider the noncommutative ring $R^{\mathrm{int}}_{\mathbb{G}_m} := \mathbb{C}[z, t^{\pm}]\langle z^2\partial_z, tz\partial_t\rangle$. A crucial point was to show that a certain quotient of the corresponding sheaf $\mathcal{R}^{\mathrm{int}}_{\mathbb{G}_m}$ on $\mathbb{G}_m$ which restricts to the $\mathcal{D}_{\mathbb{G}_{m,t}}$-module $\mathcal{D}_{\mathbb{G}_m}/\mathcal{D}_{\mathbb{G}_m}P$ on $z = 1$ actually underlies an object in the category $\mathrm{IrrMHM}(\mathbb{G}_m)$ and the latter can be uniquely extended to an object in $\mathrm{IrrMHM}(\mathbb{P}^1)$.

In this paper we discuss the case of more general hypergeometric $\mathcal{D}$-module, that is, for quotients $\mathcal{D}_{\mathbb{G}_m}/\mathcal{D}_{\mathbb{G}_m}P$, where now P is of the form

$$P = \prod_{i=1}^{n}(t\partial_t - \alpha_i) - t\prod_{j=1}^{m}(t\partial_t - \beta_j)$$

for positive integers m, n and real numbers $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_m$ such that there is no integer difference between any α_i and β_j (this is the irreducibility assumption). It is worth noticing that the presence of the factor $\prod_{j=1}^{m}(t\partial_t - \beta_j)$ rules out the usage of the geometric arguments of [Castaño Domínguez and Sevenheck 2019]. We obtain (see Theorem 5.7) that for certain such systems, the corresponding quotient of $\mathcal{R}^{\mathrm{int}}_{\mathbb{G}_m}$ still underlies an object of $\mathrm{IrrMHM}(\mathbb{G}_m)$. As an application, we can completely determine the irregular Hodge filtration for all systems $\mathcal{D}/\mathcal{D}P$ as above, where n is arbitrary and where $m = 1$.

The strategy of the proof (which is rather different from that of [Castaño Domínguez and Sevenheck 2019]) of the main theorem is to reduce these differential systems from (Fourier–Laplace transformed) A-hypergeometric $\mathcal{D}$-modules (the so-called GKZ-systems of Gelfand, Graev, Zelevinski and Kapranov, see [Gelfand et al. 1987; Gelfand et al. 1989]), but at the level of (algebraic, integrable, mixed) twistor $\mathcal{D}$-modules. Notice that the paper [Mochizuki 2015b] also studies twistor structures on GKZ-systems, by considering twistor $\mathcal{D}$-modules associated to meromorphic functions. We use instead a central result of [Reichelt and Sevenheck 2015], where the Hodge filtration on certain GKZ-systems has been computed explicitly. Technically, the main point in our proof consists in showing that for an $\mathcal{R}$-module underlying an integrable mixed twistor $\mathcal{D}$-module on the affine space, the algebraic Fourier–Laplace transformation (which is defined very much the same as in the case of algebraic $\mathcal{D}$-modules) coincides with the Fourier–Laplace transformation that can be defined inside the category MTM, or even IrrMHM. Along the way, we also obtain (see Theorem 4.7) that an $\mathcal{R}$-module version of the GKZ-$\mathcal{D}$-module underlies an irregular Hodge module provided that the parameter $\beta \in \mathbb{C}^d$ of this system satisfies a natural combinatorial condition. Notice that for the special case $\beta = 0$, this theorem can also be deduced from [Mochizuki 2015b, Proposition 1.4].

Our results give concrete representations for objects in the categories MTM and IrrMHM, which usually are difficult to describe explicitly. We hope that a similar approach can be used to understand the irregular Hodge filtration for some higher dimensional analogues of the classical hypergeometric systems, also called Horn systems, which occur in the mirror symmetry picture for toric varieties.

2. Some results on $\mathscr{R}$- and mixed twistor $\mathcal{D}$-modules

Let X be a complex manifold of dimension d. We denote by $\mathscr{O}_X$ the sheaf of holomorphic functions and $\mathscr{D}_X$ the sheaf of differential operators with holomorphic coefficients. Recall that $\mathscr{D}_X$ is generated by the tangent sheaf Θ_X. We put $\mathscr{X} := \mathbb{A}^1_z \times X$, where the subscript means that z is the canonical coordinate on $\mathbb{A}^1$. Denote by $p_z : \mathscr{X} \to X$ the projection. We denote by $\mathscr{R}_\mathscr{X}$ the sheaf of subalgebras of $\mathscr{D}_\mathscr{X}$ generated by $zp_z^*\Theta_X$ over $\mathscr{O}_\mathscr{X}$ and by $\mathscr{R}^{\text{int}}_\mathscr{X}$ the sheaf of subalgebras of $\mathscr{D}_\mathscr{X}$ generated by $zp_z^*\Theta_X$ and $z^2\partial_z$ over $\mathscr{O}_\mathscr{X}$. In local coordinates $x_1, \dots, x_d$, they are given by $\mathscr{O}_\mathscr{X}\langle z\partial_{x_1}, \dots, z\partial_{x_d}\rangle$ and $\mathscr{O}_\mathscr{X}\langle z^2\partial_z, z\partial_{x_1}, \dots, z\partial_{x_d}\rangle$, respectively. We set $\Omega^1_\mathscr{X} := z^{-1}p_z^*\Omega^1_X$ as a subsheaf of $p_z^*\Omega^1_X \otimes \mathscr{O}_\mathscr{X}(*(\{0\} \times X))$, $\Omega^p_\mathscr{X} := \bigwedge^p \Omega^1_\mathscr{X}$ and $\omega_\mathscr{X} := \Omega^d_\mathscr{X}$.

Let $f : X \to Y$ be a morphism of complex manifolds. We consider the transfer $\mathscr{R}$-modules, given by $\mathscr{R}_{\mathscr{X}\to\mathscr{Y}} := \mathscr{O}_\mathscr{X} \otimes_{f^{-1}\mathscr{O}_\mathscr{Y}} f^{-1}\mathscr{R}_\mathscr{Y}$ and $\mathscr{R}_{\mathscr{Y}\leftarrow\mathscr{X}} := \omega_\mathscr{X} \otimes \mathscr{R}_{\mathscr{X}\to\mathscr{Y}} \otimes f^{-1}\omega_\mathscr{Y}$, being respectively a $(\mathscr{R}_\mathscr{X}, f^{-1}\mathscr{R}_\mathscr{Y})$-bimodule and a $(f^{-1}\mathscr{R}_\mathscr{Y}, \mathscr{R}_\mathscr{X})$-bimodule. We have the inverse image and direct image functors

$$f^+(\mathcal{N}) := \mathscr{R}_{\mathscr{X}\to\mathscr{Y}} \overset{\boldsymbol{L}}{\otimes}_{f^{-1}\mathscr{R}_\mathscr{Y}} f^{-1}\mathcal{N}, \quad f_+(\mathcal{M}) := \boldsymbol{R} f_*(\mathscr{R}_{\mathscr{Y}\leftarrow\mathscr{X}} \overset{\boldsymbol{L}}{\otimes}_{\mathscr{R}_\mathscr{X}} \mathcal{M}), \tag{1}$$

between the bounded derived categories $D^b(\mathscr{R}_\mathscr{X})$ and $D^b(\mathscr{R}_\mathscr{Y})$.

If $f : X \times Y \to Y$ is a projection and $\dim X = d$, then $f_+(\mathcal{M})$ is given by

$$f_+(\mathcal{M}) = \boldsymbol{R} f_* \operatorname{DR}_{\mathscr{X}\times\mathscr{Y}/\mathscr{Y}}(\mathcal{M})[d],$$

where $\operatorname{DR}_{\mathscr{X}\times\mathscr{Y}/\mathscr{Y}}(\mathcal{M})$ is the relative de Rham complex with differential

$$d(\eta \otimes m) = d\eta \otimes m + \sum_{i=1}^d \left(\frac{dx_i}{z} \wedge \eta\right) \otimes z\partial_{x_i} m,$$

the $(x_i)_{1\le i\le d}$ being local coordinates on X.

Let $\sigma : \mathbb{G}_{m,z} \to \mathbb{G}_{m,z}$ be the automorphism $z \mapsto -\bar{z}^{-1}$. Set $\mathbf{S} := \{z \in \mathbb{A}^1_z \mid |z| = 1\}$. If $\lambda \in \mathbf{S}$ then $\sigma(\lambda) = -\lambda$. Let $\mathcal{E}^{(d,d)}_{\mathbf{S}\times X/\mathbf{S},c}(V)$ be the space of C^∞-sections of $\Omega^{d,d}_{\mathbf{S}\times X/\mathbf{S}}$ over any open subset V of $\mathbf{S} \times X$ with compact support and $C^0_c(\mathbf{S})$ the space of continuous functions on $\mathbf{S}$ with compact support. The space of $C^\infty(\mathbf{S})$-linear maps $\mathcal{E}^{(n,n)}_{\mathbf{S}\times X/\mathbf{S},c}(V) \to C^0_c(\mathbf{S})$ is denoted by $\mathfrak{Db}_{\mathbf{S}\times X/\mathbf{S}}(V)$. This gives rise to the sheaf $\mathfrak{Db}_{\mathbf{S}\times X/\mathbf{S}}$. The abelian category $\mathscr{R}$-Tri(X) consists of triples $\mathcal{T} = (\mathcal{M}_1, \mathcal{M}_2, C)$ where $\mathcal{M}_1, \mathcal{M}_2$ are $\mathscr{R}_\mathscr{X}$-modules and $C : \mathcal{M}_{1|\mathbf{S}\times X} \otimes \sigma^*\mathcal{M}_{2|\mathbf{S}\times X} \to \mathfrak{Db}_{\mathbf{S}\times X/\mathbf{S}}$ is a $\mathscr{R}_{\mathscr{X}|\mathbf{S}\times X} \otimes \sigma^*\mathscr{R}_{\mathscr{X}|\mathbf{S}\times\mathscr{X}}$-linear morphism. If $D \subset X$ is a hypersurface, one similarly defines a category $\mathscr{R}$-Tri(X, D) using $\mathscr{R}_\mathscr{X}(*D) := \mathscr{R}_\mathscr{X} \otimes_{\mathscr{O}_\mathscr{X}} \mathscr{O}_\mathscr{X}(*(\mathbb{A}^1_z \times D))$-modules (see [Mochizuki 2015a, §2.1] for details).

Now let $X := X_0 \times \mathbb{A}^1_t$ and let $\Theta_X(\log X_0)$ be the sheaf of vector fields on X which are logarithmic along X_0. Let $V_0\mathscr{R}_\mathscr{X}$ be the sheaf of subalgebras in $\mathscr{R}_\mathscr{X}$ which is generated by $zp_z^*\Theta_X(\log X_0)$. For $z_0 \in \mathbb{A}^1_z$ we denote by $\mathscr{X}^{(z_0)}$ a small neighborhood of $\{z_0\} \times X$. A coherent $\mathscr{R}_\mathscr{X}$-module is called strictly specializable along t at z_0 if $\mathcal{M}_{|\mathscr{X}^{(z_0)}}$ is equipped with an increasing and exhaustive filtration $V^{(z_0)}_a(\mathcal{M}_{|\mathscr{X}^{(z_0)}})_{a\in\mathbb{R}}$ by coherent $(V_0\mathscr{R}_\mathscr{X})_{|\mathscr{X}^{(z_0)}}$-modules satisfying certain conditions (see [Mochizuki 2015a, §§2.1.2.1, 2.1.2.2]). This filtration is unique if it exists. $\mathcal{M}$ is called strictly specializable along t if it is strictly specializable along t for any z_0.

Remark 2.1. If $\mathcal{M}$ is itself a coherent $V_0\mathscr{R}_\mathscr{X}$-module, then $\mathcal{M}$ is automatically specializable along t and the corresponding filtration $V_a(\mathcal{M})$ exists globally and is trivial, i.e., $V_a(\mathcal{M}) = V_b(\mathcal{M})$ for all $a, b \in \mathbb{R}$.

If $\mathcal{M}$ is a coherent $\mathscr{R}_{\mathscr{X}}(*t)$-module, we define similarly a filtration $V_a^{(z_0)}(\mathcal{M}_{|\mathscr{X}^{(z_0)}})$ and the notion of strict specializability along t (see [Mochizuki 2015a, §3.1.1]). In this case we define the $\mathscr{R}_{\mathscr{X}}$-submodules $\mathcal{M}[*t]$ and $\mathcal{M}[!t]$ of $\mathcal{M}$, which are locally generated by $V_0^{(z_0)}\mathcal{M}$ and $V_{<0}^{(z_0)}\mathcal{M}$, respectively.

Remark 2.2. If the coherent $\mathscr{R}_{\mathscr{X}}(*t)$-module $\mathcal{M}$ is itself $V_0\mathscr{R}_{\mathscr{X}}$ coherent, then

$$\mathcal{M}[!t] = \mathcal{M}[*t] = \mathcal{M}(*t) = \mathcal{M}.$$

Given an $\mathscr{R}_{\mathscr{X}}(*t)$-triple $\mathcal{T} = (\mathcal{M}_1, \mathcal{M}_2, C)$ which is strictly specializable along t we can define

$$\mathcal{T}[!t] := (\mathcal{M}_1[*t], \mathcal{M}_2[!t], C[!t]), \quad \mathcal{T}[*t] := (\mathcal{M}_1[!t], \mathcal{M}_2[*t], C[*t])$$

(see [Mochizuki 2015a, Proposition 3.2.1] for details).

The category of filtered $\mathscr{R}_{\mathscr{X}}$-triples (i.e., $\mathscr{R}_{\mathscr{X}}$-triples equipped with a finite increasing filtration W) underlies the category $\mathrm{MTM}(X)$ of mixed twistor $\mathcal{D}$-modules (see [Mochizuki 2015a, Definition 7.2.1]). The full subcategory of objects $\mathcal{T} \in \mathrm{MTM}(X)$ satisfying $\mathcal{T} = \mathcal{T}[*D]$ for some hypersurface $D \subset X$ is denoted by $\mathrm{MTM}(X, [*D])$.

If X is a smooth, algebraic variety, we denote by X^{an} the corresponding complex manifold. Let $\overline{X}$ be a smooth, complete, algebraic variety such that $X \hookrightarrow \overline{X}$ is an open immersion and $D := \overline{X} \setminus X$ is a hypersurface. We can define the category of (integrable) algebraic, mixed twistor $\mathcal{D}$-modules as

$$\mathrm{MTM}_{\mathrm{alg}}^{(\mathrm{int})}(X) := \mathrm{MTM}^{(\mathrm{int})}(\overline{X}^{\mathrm{an}}, [*D]). \tag{2}$$

We remark that this definition is independent of the completion up to an equivalence of categories [Mochizuki 2015a, Lemma 14.1.3].

Let $f : X \to Y$ be a quasiprojective morphism of smooth, algebraic varieties. We take completions $X \subset \overline{X}$ and $Y \subset \overline{Y}$ as above, such that $D_X := \overline{X} \setminus X$ and $D_Y := \overline{Y} \setminus Y$ and we have a projective morphism $\bar{f} : \overline{X} \to \overline{Y}$ which restricts to f. For $\mathcal{T} \in \mathrm{MTM}_{\mathrm{alg}}(X)$, corresponding to $\overline{\mathcal{T}} \in \mathrm{MTM}(\overline{X}, [*D_X])$, we define

$$f_*^i \mathcal{T} := \mathcal{H}^i \bar{f}_* \overline{\mathcal{T}},$$

where $\bar{f}_*$ is the direct image functor for mixed twistor $\mathcal{D}$-modules arising from the one for $\mathscr{R}$-modules depicted in (1).

If X is an algebraic variety, we denote by $\mathcal{D}_X$ the sheaf of algebraic differential operators and by $\mathcal{R}_{\mathscr{X}}$ the sheaf of z-differential operators, where here $\mathscr{X} := \mathbb{A}_z^1 \times X$. We define the inverse and direct image functor in the category of algebraic $\mathcal{R}_{\mathscr{X}}$-modules as in (1). Analogously to the construction of $\mathcal{R}_{\mathscr{X}}$, we can consider the projection $p : \mathbb{P}^1 \times X \to X$, and construct the sheaf of subalgebras of $\mathcal{D}_{\mathbb{P}^1 \times X}(*(\{\infty\} \times X))$ generated by $z^2\partial_z$ and $zp^*\Theta_X$ over $\mathcal{O}_{\mathbb{P}^1 \times X}$ (see [Mochizuki 2015a, §14.4.1.1]), which will be denoted by $\mathcal{R}_{\mathbb{P}^1 \times X}^{\mathrm{int}}(*\infty)$. In that sense, an algebraic integrable $\mathcal{R}_{\mathscr{X}}$-module gives rise to a unique $\mathcal{R}_{\mathbb{P}^1 \times X}^{\mathrm{int}}(*\infty)$-module (see [ibid., Theorem 14.4.8]).

The following lemma, which will be needed later, is due to T. Mochizuki.

Lemma 2.3. *Given two good $\mathcal{R}_{\mathbb{P}^1 \times X}^{\mathrm{int}}(*\infty)$-modules $\mathcal{P}_1, \mathcal{P}_2$ and an analytic isomorphism $f : \mathcal{P}_1^{\mathrm{an}} \to \mathcal{P}_2^{\mathrm{an}}$, then f is induced by a unique algebraic isomorphism between $\mathcal{P}_1$ and $\mathcal{P}_2$.*

Proof. Take a coherent $\mathcal{O}_{\mathbb{P}^1\times X}$-submodule $\mathcal{N}_1 \subset \mathcal{P}_1$ such that $\mathcal{R}^{\text{int}}_{\mathbb{P}^1\times X}(*\infty)\otimes\mathcal{N}_1 \to \mathcal{P}_1$ is surjective and a coherent $\mathcal{O}_{\mathbb{P}^1\times X}$-module $\mathcal{N}_2 \subset \mathcal{P}_2$ such that both $\mathcal{R}^{\text{int}}_{\mathbb{P}^1\times X}(*\infty)\otimes\mathcal{N}_2 \to \mathcal{P}_2$ is surjective and $f(\mathcal{N}_1^{\text{an}}) \subset \mathcal{N}_2^{\text{an}}$. According to GAGA we have a morphism $g : \mathcal{N}_1 \to \mathcal{N}_2$ which after analytification is equal to the morphism $\mathcal{N}_1^{\text{an}} \to \mathcal{N}_2^{\text{an}}$ induced by f. Denote by $\mathcal{K}_1$ the kernel of $\mathcal{R}^{\text{int}}_{\mathbb{P}^1\times X}(*\infty)\otimes\mathcal{N}_1 \to \mathcal{P}_1$. This gives a morphism $\mathcal{K}_1 \to \mathcal{P}_2$ which one obtains as the composition $\mathcal{K}_1 \to \mathcal{R}^{\text{int}}_{\mathbb{P}^1\times X}(*\infty)\otimes\mathcal{N}_1 \xrightarrow{\varphi} \mathcal{R}^{\text{int}}_{\mathbb{P}_1\times}(*\infty)\otimes\mathcal{N}_2 \to \mathcal{P}_2$, where φ is induced by g. Because the induced morphism $(\mathcal{R}^{\text{int}}_{\mathbb{P}^1\times X}(*\infty)\otimes\mathcal{N}_1)^{\text{an}} \to \mathcal{P}_2^{\text{an}}$ factors through $\mathcal{P}_1^{\text{an}}$, the induced morphism $\mathcal{K}_1^{\text{an}} \to \mathcal{P}_2^{\text{an}}$ is 0. Hence, we obtain that $\mathcal{K}_1 \to \mathcal{P}_2$ is 0, which means that $\mathcal{R}^{\text{int}}_{\mathbb{P}^1\times X}(*\infty)\otimes\mathcal{N}_1 \to \mathcal{P}_2$ factors through $\mathcal{P}_1$. This shows the existence. The uniqueness follows from [Serre 1955–1956, Proposition 10]. □

Since an algebraic, integrable, mixed twistor $\mathcal{D}$-module on X gives rise to an analytic $\mathscr{R}^{\text{int}}_{\mathbb{P}^1\times X}(*\infty)$-module which underlies an algebraic $\mathcal{R}^{\text{int}}_{\mathbb{P}^1\times X}(*\infty)$-module by [Mochizuki 2015a, Theorem 14.4.8], the lemma above shows that we can define functors (up to canonical isomorphism)

$$\begin{aligned} \mathrm{For}_i : \mathrm{MTM}^{\text{int}}_{\text{alg}}(X) &\to \mathrm{Mod}(\mathcal{R}^{\text{int}}_{\mathscr{X}}) \\ (\mathcal{M}_1, \mathcal{M}_2, C) &\mapsto \mathcal{M}_i \qquad \text{for } i = 1, 2, \end{aligned}$$

which become faithful if we impose goodness.

3. Fourier transformation of twistor $\mathcal{D}$-modules

In this section we define the Fourier–Laplace transformation in the categories of integrable $\mathcal{R}$-modules and integrable, algebraic, mixed twistor $\mathcal{D}$-modules, and we prove that these two transformations are compatible.

Consider the diagram

$$\begin{array}{ccccc} & & \mathbb{A}^N\times\hat{\mathbb{A}}^N & \xrightarrow{\quad j \quad} & \mathbb{P}^N\times\hat{\mathbb{P}}^N\ , \\ & {\scriptstyle p}\swarrow & & \searrow{\scriptstyle q} & \downarrow{\scriptstyle \bar{q}} \\ \mathbb{A}^N & & & \hat{\mathbb{A}}^N \xrightarrow{\ \hat{j}\ } & \hat{\mathbb{P}}^N \end{array}$$

where p and q are the projections to the first and second factor respectively. Consider the function $\varphi = \sum_{i=1}^N w_i \cdot \lambda_i$ on $\mathbb{A}^N \times \hat{\mathbb{A}}^N$.

Let $\mathcal{A}^{\varphi/z}_{\text{aff}}$ be the $\mathcal{R}_{\mathbb{A}^1\times\mathbb{A}^N\times\hat{\mathbb{A}}^N}$-module $\mathcal{O}_{\mathbb{A}^1\times\mathbb{A}^N\times\hat{\mathbb{A}}^N}$ equipped with the z-connection $zd + d\varphi$, and consider the reduced divisor $D := (\mathbb{P}^N \times \hat{\mathbb{P}}^N) \setminus (\mathbb{A}^N \times \hat{\mathbb{A}}^N)$. Then $\mathcal{A}^{\varphi/z}_* := j_*\mathcal{A}^{\varphi/z}_{\text{aff}}$ carries a natural structure of an $\mathcal{R}_{\mathbb{A}^1\times\mathbb{P}^N\times\hat{\mathbb{P}}^N}(*D)$-module.

We denote by $\mathcal{E}^{\varphi/z}_*$ the analytification of $\mathcal{A}^{\varphi/z}_*$, which is an $\mathscr{R}_{\mathbb{A}^1\times\mathbb{P}^N\times\hat{\mathbb{P}}^N}(*D)$-module.

Lemma 3.1. *$\mathcal{E}^{\varphi/z}_*$ is strictly specializable along D and*

$$\mathcal{E}^{\varphi/z} := \mathcal{E}^{\varphi/z}_*[*D] = \mathcal{E}^{\varphi/z}_*.$$

Proof. We denote the coordinates on $\mathbb{P}^N \times \hat{\mathbb{P}}^N$ by $((w_0 : w_1 : \cdots : w_N), (\lambda_0 : \lambda_1 : \cdots : \lambda_N))$, where the chart $\mathbb{A}^N \times \hat{\mathbb{A}}^N$ is embedded via the map $j : (w_1, \ldots, w_N, \lambda_1, \ldots, \lambda_N) \mapsto ((1 : w_1, \ldots, w_N), (1 : \lambda_1 : \cdots : \lambda_N))$. By symmetry it is enough to prove the claim in the charts $\{w_1 \neq 0, \lambda_0 \neq 0\}$, $\{w_1 \neq 0, \lambda_1 \neq 0\}$ and $\{w_1 \neq 0, \lambda_2 \neq 0\}$. We will assume $N \geq 2$ and consider the chart $X := \{w_1 \neq 0, \lambda_2 \neq 0\}$; the arguments with the other charts and when $N = 1$ go similarly. The chart X is embedded as $(x_1, \ldots, x_N, \mu_1, \ldots, \mu_N) \mapsto ((x_1 : 1 : x_2 \cdots : x_N), (\mu_1 : \mu_2 : 1 : \mu_3 : \cdots : \mu_N))$, so that the map φ is given on X by

$$\frac{1}{x_1\mu_1}\Big(\mu_2 + x_2 + \sum_{i\geq 3} \mu_i x_i\Big).$$

Set $\mathcal{D}_X := \mathbb{A}^1 \times (D \cap X) = \mathbb{A}^1 \times \{x_1 \cdot \mu_1 = 0\}$. The module $(\mathcal{E}_*^{\varphi/z})_{|X}$ is a cyclic $\mathcal{R}_{\mathbb{A}^1\times X}(*\mathcal{D}_X)$-module $\mathcal{R}_{\mathbb{A}^1\times X}(*\mathcal{D}_X)/\mathcal{I}$, where the left ideal $\mathcal{I}$ is generated by

$$z\partial_{x_1} + \frac{1}{x_1^2\mu_1}\Big(\mu_2 + x_2 + \sum_{i\geq 3}\mu_i x_i\Big), \quad z\partial_{x_2} - \frac{1}{x_1\mu_1}, \quad z\partial_{x_j} - \frac{\mu_j}{x_1\mu_1},$$
$$z\partial_{\mu_1} + \frac{1}{x_1\mu_1^2}\Big(\mu_2 + x_2 + \sum_{i\geq 3}\mu_i x_i\Big), \quad z\partial_{\mu_2} - \frac{1}{x_1\mu_1}, \quad z\mu_j - \frac{x_j}{x_1\mu_1},$$

where $j \geq 3$. Consider the map $i_g : X \to \mathbb{A}^1_t \times X$ given by

$$(x_1, \ldots, x_N, \mu_1, \ldots, \mu_N) \mapsto (x_1 \cdot \mu_1, x_1, \ldots, x_N, \mu_1, \ldots, \mu_N).$$

The direct image $i_{g,+}(\mathcal{R}_\mathcal{X}(*\mathcal{D}_X)/\mathcal{I})$ is a cyclic $\mathcal{R}_{\mathbb{A}^1_t\times\mathcal{X}}(*(\mathbb{A}^1_t \times \mathcal{D}_X))$-module $\mathcal{R}_{\mathbb{A}^1_t\times\mathcal{X}}(*(\mathbb{A}^1_t \times \mathcal{D}_X))/\mathcal{J}'$ where $\mathcal{J}'$ is generated by

$$z\partial_{x_1} + \mu_1 z\partial_t + \frac{1}{x_1^2\mu_1}\Big(\mu_2 + x_2 + \sum_{i\geq 3}\mu_i x_i\Big), \quad z\partial_{x_2} - \frac{1}{x_1\mu_1}, \quad z\partial_{x_j} - \frac{\mu_j}{x_1\mu_1},$$
$$z\partial_{\mu_1} + x_1 z\partial_t + \frac{1}{x_1\mu_1^2}\Big(\mu_2 + x_2 + \sum_{i\geq 3}\mu_i x_i\Big), \quad z\partial_{\mu_2} - \frac{1}{x_1\mu_1}, \quad z\partial_{\mu_j} - \frac{x_j}{x_1\mu_1}, \quad t - x_1\mu_1,$$

where $j \geq 3$. Define the cyclic $\mathcal{R}_{\mathbb{A}^1_t\times\mathcal{X}}(*t)$-module $\mathcal{R}_{\mathbb{A}^1_t\times\mathcal{X}}(*t)/\mathcal{J}$ where $\mathcal{J}$ is generated by

$$z\partial_{x_1} + \mu_1 z\partial_t + \frac{\mu_1}{t^2}\Big(\mu_2 + x_2 + \sum_{i\geq 3}\mu_i x_i\Big), \quad z\partial_{x_2} - \frac{1}{t}, \quad z\partial_{x_j} - \frac{\mu_j}{t},$$
$$z\partial_{\mu_1} + x_1 z\partial_t + \frac{x_1}{t^2}\Big(\mu_2 + x_2 + \sum_{i\geq 3}\mu_i x_i\Big), \quad z\partial_{\mu_2} - \frac{1}{t}, \quad z\partial_{\mu_j} - \frac{x_j}{t}, \quad t - x_1\mu_1,$$

where $j \geq 3$. Then we have the following $\mathcal{R}_{\mathbb{A}^1_t\times\mathcal{X}}$-linear isomorphism

$$\mathcal{R}_{\mathbb{A}^1_t\times\mathcal{X}}(*(\mathbb{A}^1_t \times \mathcal{D}_X))/\mathcal{J}' \to \mathcal{R}_{\mathbb{A}^1_t\times\mathcal{X}}(*t)/\mathcal{J}$$
$$P \cdot \frac{1}{(x_1\mu_1)^k} \mapsto P \cdot \frac{1}{t^k}.$$

Consider the V-filtration along $t=0$. The relations $1/t^k=(z\partial_{\mu_2})^k$,

$$z\partial_t=-\tfrac{1}{t}\left(z\partial_{x_1}x_1+\tfrac{1}{t}\left(\mu_2+x_2+\sum_{i\geq 3}\mu_i x_i\right)\right)=-z\partial_{x_1}x_1 z\partial_{\mu_2}-\left(\mu_2+x_2+\sum_{i\geq 3}\mu_i x_i\right)(z\partial_{\mu_2})^2$$

and a straightforward induction over k for $(z\partial_t)^k$ show that $i_{g,+}(\mathscr{R}_{\mathscr{X}}(*\mathscr{D}_X)/\mathscr{J})$ is a cyclic, hence also coherent, $V_0\mathscr{R}_{\mathbb{A}^1_s\times\mathscr{X}}$-module. It follows from Remark 2.1 that $i_{g,+}(\mathscr{R}_{\mathscr{X}}(*\mathscr{D}_X)/\mathscr{J})=i_{g,+}(\mathscr{R}_{\mathscr{X}}(*\mathscr{D}_X)/\mathscr{J})[*t]$, and as a consequence, we are done by applying [Mochizuki 2015a, §3.3.1.1] and Remark 2.2. □

It follows from [Sabbah and Yu 2015, Proposition 3.3] that $\mathcal{E}^{\varphi/z}$ underlies an object $\mathcal{T}^{\varphi/z}\in\mathrm{MTM}^{\mathrm{int}}_{\mathrm{alg}}(\mathbb{A}^N\times\hat{\mathbb{A}}^N)$. Let us notice that the preceding lemma, as well as the similar Lemma 3.6 below, are related to a more general statement in [Mochizuki 2015b, Corollary 3.12] on mixed twistor $\mathcal{D}$-modules associated to nondegenerate functions. However, in order to keep the paper self-contained, we prefer to give direct proofs here.

We will now define a Fourier–Laplace transformation for algebraic $\mathcal{R}^{\mathrm{int}}_{\mathbb{A}^1\times\mathbb{A}^N}$-modules.

Definition 3.2. The Fourier–Laplace transformation functor from the category of algebraic $\mathcal{R}^{\mathrm{int}}_{\mathbb{A}^1\times\mathbb{A}^N}$-modules to the category of algebraic $\mathcal{R}^{\mathrm{int}}_{\mathbb{A}^1\times\hat{\mathbb{A}}^N}$-modules is defined as

$$\widehat{\mathcal{M}}:=\mathrm{FL}(\mathcal{M}):=\mathcal{H}^0q_+((p^+\mathcal{M})\otimes\mathcal{A}^{\varphi/z}_{\mathrm{aff}}),$$

for any $\mathcal{M}$ in $\mathrm{Mod}(\mathcal{R}^{\mathrm{int}}_{\mathbb{A}^1\times\mathbb{A}^N})$.

Remark 3.3. Let $M:=\Gamma(\mathbb{A}^1\times\mathbb{A}^N,\mathcal{M})$ be the $R^{\mathrm{int}}_{\mathbb{A}^1\times\mathbb{A}^N}$-module of global sections of $\mathcal{M}$. The $R^{\mathrm{int}}_{\mathbb{A}^1\times\hat{\mathbb{A}}^N}$-module $\widehat{M}:=\Gamma(\mathbb{A}^1\times\hat{\mathbb{A}}^N,\widehat{\mathcal{M}})$ is isomorphic to M as a $\mathbb{C}[z]$-module and the full $R^{\mathrm{int}}_{\mathbb{A}^1\times\hat{\mathbb{A}}^N}$-structure is given by

$$\lambda_i\cdot m:=-z\partial_{w_i}\cdot m,\quad z\partial_{\lambda_i}\cdot m:=w_i\cdot m\quad\text{and}\quad z^2\partial_z\cdot m:=\left(z^2\partial_z-\sum_{i=1}^N z\partial_{w_i}w_i\right)\cdot m.$$

On the other hand, there is a similar definition of a Fourier–Laplace transformation in the category of algebraic $\mathcal{D}_{\mathbb{A}^N}$-modules (see e.g., [Reichelt 2014, Definition 1.2]) which we also denote by FL.

The Fourier–Laplace transformation for algebraic, integrable, mixed twistor $\mathcal{D}$-modules is defined in the following way.

Definition 3.4. The Fourier–Laplace transformation in the category of algebraic, integrable mixed twistor $\mathcal{D}$-modules on $\mathbb{A}^N$ is defined by

$$\mathrm{FL}_{\mathrm{MTM}}(\mathcal{M}):=\mathcal{H}^0q_*((p^*\mathcal{M})\otimes\mathcal{T}^{\varphi/z}),$$

where $\mathcal{M}\in\mathrm{MTM}^{\mathrm{int}}_{\mathrm{alg}}(\mathbb{A}^N)$.

Recall that for $\mathcal{M}=(\mathcal{M}_1,\mathcal{M}_2,C)\in\mathrm{MTM}^{\mathrm{int}}_{\mathrm{alg}}(X)$ we denote by For_i the forgetful functors $\mathrm{For}_i(\mathcal{M})=\mathcal{M}_i$ for $i=1,2$.

Proposition 3.5. *Let $\mathcal{M} \in \mathrm{MTM}^{\mathrm{int}}_{\mathrm{alg}}(\mathbb{A}^N)$. Then*

$$\mathrm{For}_1(\mathrm{FL}_{\mathrm{MTM}}(\mathcal{M})) = \mathrm{FL}(\mathrm{For}_1(\mathcal{M})) \quad \textit{and} \quad \mathrm{For}_2(\mathrm{FL}_{\mathrm{MTM}}(\mathcal{M})) = z^{-N}\,\mathrm{FL}(\mathrm{For}_2(\mathcal{M})).$$

Proof. By [Mochizuki 2015a, §14.3.3.3] it is clear that For_i almost commutes with p^*, more precisely we have

$$\mathrm{For}_1(p^*(\mathcal{M})) = z^N p^+(\mathrm{For}_1(\mathcal{M})) \quad \text{and} \quad \mathrm{For}_2(p^*(\mathcal{M})) = p^+(\mathrm{For}_2(\mathcal{M})).$$

Then it is enough to prove for $\mathcal{N} \in \mathrm{MTM}^{\mathrm{int}}_{\mathrm{alg}}(\mathbb{A}^N \times \hat{\mathbb{A}}^N)$ that $q_+(\mathrm{For}_i(\mathcal{N}) \otimes \mathcal{A}^{\varphi/z}_{\mathrm{aff}}) \cong \mathrm{For}_i(q_*(\mathcal{N} \otimes \mathcal{T}^{\varphi/z}))$. We have

$$\begin{aligned}\hat{j}_+ q_+(\mathrm{For}_i(\mathcal{N}) \otimes \mathcal{A}^{\varphi/z}_{\mathrm{aff}}) &\cong \bar{q}_+ j_+(\mathrm{For}_i(\mathcal{N}) \otimes \mathcal{A}^{\varphi/z}_{\mathrm{aff}}) \\ &\cong \bar{q}_+ j_*(\mathrm{For}_i(\mathcal{N}) \otimes \mathcal{A}^{\varphi/z}_{\mathrm{aff}}) \\ &\cong \boldsymbol{R}\bar{q}_*\,\mathrm{DR}_{\mathbb{P}^N \times \hat{\mathbb{P}}^N}\, j_*(\mathrm{For}_i(\mathcal{N}) \otimes \mathcal{A}^{\varphi/z}_{\mathrm{aff}}).\end{aligned}$$

Since $\mathcal{N}, \mathcal{T}^{\varphi/z} \in \mathrm{MTM}^{\mathrm{int}}_{\mathrm{alg}}(\mathbb{A}^N \times \hat{\mathbb{A}}^N)$, there exist mixed twistor $\mathcal{D}$-modules $\overline{\mathcal{N}}, \overline{\mathcal{T}}^{\varphi/z} \in \mathrm{MTM}^{\mathrm{int}}(\mathbb{P}^N \times \hat{\mathbb{P}}^N, [*D])$ whose underlying $\mathscr{R}$-modules are (after stupid localization along D) analytifications of the $j_*\,\mathrm{For}_i(\mathcal{N})$ and $j_*\mathcal{A}^{\varphi/z}_{\mathrm{aff}}$. Hence

$$(j_*(\mathrm{For}_i(\mathcal{N}) \otimes \mathcal{A}^{\varphi/z}_{\mathrm{aff}}))^{\mathrm{an}} \cong \mathrm{For}_i(\overline{\mathcal{N}} \otimes \overline{\mathcal{T}}^{\varphi/z})(*D) \cong \mathrm{For}_i(\overline{\mathcal{N}} \otimes \overline{\mathcal{T}}^{\varphi/z}),$$

where the last equation follows from Lemma 3.1. We therefore get

$$\begin{aligned}(\hat{j}_+ p_+(\mathrm{For}_i(\mathcal{N}) \otimes \mathcal{A}^{\varphi/z}_{\mathrm{aff}}))^{\mathrm{an}} &\cong \boldsymbol{R}\bar{q}_*\,\mathrm{DR}^{\mathrm{an}}_{\mathbb{P}^N \times \hat{\mathbb{P}}^N}(j_*(\mathrm{For}_i(\mathcal{N}) \otimes \mathcal{A}^{\varphi/z}_{\mathrm{aff}}))^{\mathrm{an}} \\ &\cong \boldsymbol{R}\bar{q}_*\,\mathrm{DR}^{\mathrm{an}}_{\mathbb{P}^N \times \hat{\mathbb{P}}^N}\,\mathrm{For}_i(\overline{\mathcal{N}} \otimes \overline{\mathcal{T}}^{\varphi/z}) \\ &\cong \mathrm{For}_i(\bar{q}_*(\overline{\mathcal{N}} \otimes \overline{\mathcal{T}}^{\varphi/z})).\end{aligned}$$

The claim follows now from Lemma 2.3, noting that the goodness is a consequence of Lemma 3.1 and [Mochizuki 2015a, Lemma 14.4.15]. □

We have the following variant, which will be used in the next section. Consider the diagram

$$\begin{array}{ccccc} & & \mathbb{A}^N \times \mathbb{G}_m & \xrightarrow{\quad j \quad} & \mathbb{P}^N \times \mathbb{P}^1 \\ & \swarrow^{p} & & \searrow^{q} & \downarrow \bar{q} \\ \mathbb{A}^N & & & & \mathbb{G}_m \xrightarrow{\ \hat{j}\ } \mathbb{P}^1 \end{array}$$

and let $\psi := w_1 \cdot t + w_2 + \cdots + w_N$.

As above we define the $\mathcal{R}_{\mathbb{A}^1 \times \mathbb{A}^N \times \mathbb{G}_m}$-module $\mathcal{A}^{\psi/z}_{\mathrm{aff}}$, being $\mathcal{O}_{\mathbb{A}^1 \times \mathbb{A}^N \times \mathbb{G}_m}$ endowed with the z-connection $zd + d\psi$. As in the other case, we can consider the divisor $H := (\mathbb{P}^N \times \mathbb{P}^1) \setminus (\mathbb{A}^N \times \mathbb{G}_m)$ and obtain the $\mathcal{R}_{\mathbb{A}^1 \times \mathbb{P}^N \times \mathbb{P}^1}(*H)$-module $\mathcal{A}^{\psi/z}_* := j_*\mathcal{A}^{\psi/z}_{\mathrm{aff}}$. In the same vein as before, we will denote by $\mathcal{E}^{\psi/z}_*$ the $\mathscr{R}_{\mathbb{A}^1 \times \mathbb{P}^N \times \mathbb{P}^1}(*H)$-module being the analytification of $\mathcal{A}^{\psi/z}_*$. The following lemma is similar to Lemma 3.1.

Lemma 3.6. *$\mathcal{E}^{\psi/z}_*$ is strictly specializable along H and*

$$\mathcal{E}^{\psi/z} := \mathcal{E}^{\psi/z}_*[*H] = \mathcal{E}^{\psi/z}_*.$$

Proof. We denote the coordinates on $\mathbb{P}^N \times \mathbb{P}^1$ by $((w_0 : w_1 : \cdots : w_N), (u : t))$, where the chart $\mathbb{A}^N \times \mathbb{G}_m$ is embedded via the map $j : (w_1, \ldots, w_N, t) \mapsto ((1 : w_1 : \cdots : w_N), (1 : t))$. We will assume $N \geq 3$ and consider the chart $X := \{w_2 \neq 0, u \neq 0\}$; the other charts behave similarly, as do the cases $N = 1, 2$. The chart X is embedded as $(x_1, \ldots, x_N, u) \mapsto ((x_1 : x_2 : 1 : x_3 : \cdots : x_N), (u : 1))$. On this chart the map ψ is given by $\frac{1}{x_1}\left(\frac{x_2}{u} + 1 + x_3 + \cdots + x_N\right)$. Set $\mathcal{H}_X := \mathbb{A}^1 \times (H \cap X) = \mathbb{A}^1 \times \{x_1 \cdot u = 0\}$. The module $(\mathcal{E}_*^{\psi/z})_{|X}$ is a cyclic $\mathcal{R}_{\mathcal{X}}(*\mathcal{H}_X)$-module $\mathcal{R}_{\mathcal{X}}(*\mathcal{H}_X)/\mathcal{I}$, where the left ideal $\mathcal{I}$ is generated by

$$z\partial_{x_1} + \frac{1}{x_1^2}\left(\frac{x_2}{u} + 1 + x_3 + \cdots + x_N\right), \quad z\partial_{x_2} - \frac{1}{x_1 u}, \quad z\partial_{x_j} - \frac{1}{x_1}, \quad z\partial_u + \frac{x_2}{x_1 u^2},$$

with $j \geq 3$. Consider the map $i_g : X \to \mathbb{A}_s^1 \times X$ given by

$$(x_1, \ldots, x_N, u) \mapsto (x_1 \cdot u, x_1, \ldots, x_N, u).$$

Analogous to Lemma 3.1, the direct image $i_{g,+}(\mathcal{R}_{\mathcal{X}}(*H_X)/\mathcal{I})$ is a cyclic $\mathcal{R}_{\mathbb{A}_s^1 \times \mathcal{X}}(*(\mathbb{A}_s^1 \times \mathcal{H}_X))$-module $\mathcal{R}_{\mathbb{A}_s^1 \times \mathcal{X}}(*(\mathbb{A}_s^1 \times \mathcal{H}_X))/\mathcal{I}'$ where $\mathcal{I}'$ is the left ideal generated by

$$z\partial_{x_1} + uz\partial_s + \frac{1}{x_1^2}\left(\frac{x_2}{u} + 1 + x_3 + \cdots + x_N\right), \quad z\partial_{x_2} - \frac{1}{x_1 u}, \quad z\partial_{x_j} - \frac{1}{x_1}, \quad z\partial_u + x_1 z\partial_s + \frac{x_2}{x_1 u^2}, \quad s - x_1 u,$$

and $j \geq 3$. Define the cyclic $\mathcal{R}_{\mathbb{A}_s^1 \times \mathcal{X}}(*s)$-module $\mathcal{R}_{\mathbb{A}_s^1 \times \mathcal{X}}(*s)/\mathcal{J}$ where $\mathcal{J}$ is generated by

$$z\partial_{x_1} + uz\partial_s + \frac{1}{s^2}(x_2 u + u^2 + x_3 u^2 + \cdots + x_N u^2), \quad z\partial_{x_2} - \frac{1}{s}, \quad z\partial_{x_j} - \frac{u}{s}, \quad z\partial_u + x_1 z\partial_s + \frac{x_1 x_2}{s^2}, \quad s - x_1 u,$$

and where $j \geq 3$. We have the following $\mathcal{R}_{\mathbb{A}_s^1 \times \mathcal{X}}$-linear isomorphism

$$\mathcal{R}_{\mathbb{A}_s^1 \times \mathcal{X}}(*(\mathbb{A}_s^1 \times \mathcal{H}_X))/\mathcal{I}' \to \mathcal{R}_{\mathbb{A}_s^1 \times \mathcal{X}}(*s)/\mathcal{J}$$
$$P\frac{1}{(x_1 u)^k} \mapsto P\frac{1}{s^k}.$$

Consider the V-filtration along $s = 0$. The relations $1/s^k = (z\partial_{x_2})^k$,

$$z\partial_s = -\frac{1}{s}\left(z + uz\partial_u + \frac{x_2}{s}\right) = -z \cdot z\partial_{x_2} - uz\partial_u z\partial_{x_2} - x_2(z\partial_{x_2})^2$$

and a straightforward induction over k for $(z\partial_s)^k$ show that $i_{g,+}(\mathcal{R}_{\mathcal{X}}(*D_X)/\mathcal{I})$ is a coherent $V_0\mathcal{R}_{\mathbb{A}_s^1 \times \mathcal{X}}$-module. As in the previous lemma, this shows the claim. □

It follows again from [Sabbah and Yu 2015, Proposition 3.3] that $\mathcal{E}^{\psi/z}$ underlies an object $\mathcal{T}^{\psi/z} \in \mathrm{MTM}_{\mathrm{alg}}^{\mathrm{int}}(\mathbb{A}^n \times \mathbb{G}_m)$.

Definition 3.7. (1) The Fourier–Laplace transformation with respect to the kernel ψ in the category of algebraic $\mathcal{R}_{\mathbb{A}^1 \times \mathbb{A}^N}$-modules is defined as

$$\mathrm{FL}^{\psi}(\mathcal{M}) := \mathcal{H}^0 q_+((p^+\mathcal{M}) \otimes \mathcal{A}_{\mathrm{aff}}^{\psi/z}),$$

for any $\mathcal{M} \in \mathrm{Mod}(\mathcal{R}_{\mathbb{A}^N})$.

(2) Analogously, the Fourier–Laplace transformation with respect to the kernel ψ in the category of algebraic, integrable twistor $\mathcal{D}$-modules on $\mathbb{A}^N$ is defined by

$$\mathrm{FL}^{\psi}_{\mathrm{MTM}}(\mathcal{M}) := \mathcal{H}^0 q_*((p^*\mathcal{M}) \otimes \mathcal{T}^{\psi/z}),$$

for any $\mathcal{M} \in \mathrm{MTM}^{\mathrm{int}}_{\mathrm{alg}}(\mathbb{A}^N)$.

We get the following result for the kernel ψ.

Proposition 3.8. *Let $\mathcal{M} \in \mathrm{MTM}^{\mathrm{int}}_{\mathrm{alg}}(\mathbb{A}^N)$. Then*

$$\mathrm{For}_1(\mathrm{FL}^{\psi}_{\mathrm{MTM}}(\mathcal{M})) = z^{1-N}\,\mathrm{FL}^{\psi}(\mathrm{For}_1(\mathcal{M})) \quad \textit{and} \quad \mathrm{For}_2(\mathrm{FL}^{\psi}_{\mathrm{MTM}}(\mathcal{M})) = z^{-N}\,\mathrm{FL}^{\psi}(\mathrm{For}_2(\mathcal{M})).$$

Proof. We have, by [Mochizuki 2015a, §14.3.3.3],

$$\mathrm{For}_1(p^*(\mathcal{M})) = zp^+(\mathrm{For}_1(\mathcal{M})) \quad \text{and} \quad \mathrm{For}_2(p^*(\mathcal{M})) = p^+(\mathrm{For}_2(\mathcal{M})).$$

The rest of the proof carries over almost word for word from Proposition 3.5, using Lemma 3.6. □

4. GKZ systems and irregular Hodge modules

Let $A = (a_{ki})$ be a $d \times N$ integer matrix with columns $(\underline{a}_1, \ldots, \underline{a}_N)$. We define

$$\mathbb{N}A := \sum_{i=1}^{N} \mathbb{N}\underline{a}_i \subset \mathbb{Z}^d$$

and similarly for $\mathbb{Z}A$ and $\mathbb{R}_{\geq 0}A$. Throughout this section we assume

$$\mathbb{Z}A = \mathbb{Z}^d \quad \text{and} \quad \mathbb{N}A = \mathbb{Z}^d \cap \mathbb{R}_{\geq 0}A.$$

Set $\mathbb{A}^N := \mathrm{Spec}(\mathbb{C}[w_1, \ldots, w_N])$ and $\hat{\mathbb{A}}^N := \mathrm{Spec}(\mathbb{C}[\lambda_1, \ldots, \lambda_N])$ and

$$\mathbb{L}_A := \left\{ \ell = (\ell_1, \ldots, \ell_N) \in \mathbb{Z}^N : \sum_{i=1}^{N} \ell_i \underline{a}_i \right\}.$$

Definition 4.1. The GKZ-hypergeometric system $\mathcal{M}^{\beta}_A$ is the cyclic $\mathcal{D}_{\hat{\mathbb{A}}^N}$-module $\mathcal{D}_{\hat{\mathbb{A}}^N}/\mathcal{I}$, where $\mathcal{I}$ is the left ideal generated by

$$E_k := \sum_{i=1}^{N} a_{ki}\lambda_i\partial_{\lambda_i} - \beta_k, \quad \text{for } k = 1, \ldots, d,$$

and

$$\Box_\ell := \prod_{\ell_i > 0} \partial_{\lambda_i}^{\ell_i} - \prod_{\ell_i < 0} \partial_{\lambda_i}^{-\ell_i}, \quad \text{for } l \in \mathbb{L}_A.$$

The GKZ-hypergeometric system $\mathcal{M}_A^\beta$ is the Fourier–Laplace transform of the cyclic $\mathcal{D}_{\mathbb{A}^N}$-module $\check{\mathcal{M}}_A^\beta := \mathcal{D}_{\mathbb{A}^N}/\mathcal{J}$, where $\mathcal{J}$ is the left ideal generated by

$$\check{E}_k := \sum_{i=1}^{N} a_{ki}\partial_{w_i} w_i + \beta_k, \quad \text{for } k = 1, \dots, d,$$

and

$$\check{\Box}_\ell := \prod_{\ell_i>0} w_i^{\ell_i} - \prod_{\ell_i<0} w_i^{-\ell_i}, \quad \text{for } l \in \mathbb{L}_A.$$

The semigroup ring $\mathbb{C}[\mathbb{N}A] \subset \mathbb{C}[t_1^\pm, \dots, t_d^\pm]$ is naturally a $\mathbb{C}[w_1, \dots, w_N]$-module under the isomorphism

$$\begin{aligned} \mathbb{C}[w_1, \dots, w_N]/((\check{\Box}_\ell)_{\ell\in\mathbb{L}_A}) &\to \mathbb{C}[\mathbb{N}A] \\ w_i &\mapsto t^{\underline{a}_i}, \end{aligned}$$

where we are using the multiindex notation $t^{\underline{a}_i} := \prod_{k=1}^d t_k^{a_{ki}}$. We set $S_A := \mathbb{C}[\mathbb{N}A]$. Notice that the rings $\mathbb{C}[w_1, \dots, w_N]$ and S_A carry a natural $\mathbb{Z}^d$-grading given by $\deg(w_i) = \underline{a}_i$. This is compatible with the grading on the Weyl algebra $D_{\mathbb{A}^N} := \Gamma(\mathbb{A}^N, \mathcal{D}_{\mathbb{A}^N})$ given by $\deg(w_i) = \underline{a}_i$ and $\deg(\partial_{w_i}) = -\underline{a}_i$.

Definition 4.2 [Matusevich et al. 2005, Definition 5.2]. Let P be a finitely generated $\mathbb{Z}^d$-graded $\mathbb{C}[w_1, \dots, w_N]$-module. An element $\alpha \in \mathbb{Z}^d$ is called a true degree of P if the graded part P_α is nonzero. A vector $\alpha \in \mathbb{C}^d$ is called a quasidegree of P if α lies in the complex Zariski closure $\operatorname{qdeg}(P)$ of the true degrees of P via the natural embedding $\mathbb{Z}^d \hookrightarrow \mathbb{C}^d$.

Consider the set of strongly resonant parameters of A:

$$\operatorname{sRes}(A) := \bigcup_{j=1}^{N} \operatorname{sRes}_j(A),$$

where

$$\operatorname{sRes}_j(A) := \{\beta \in \mathbb{C}^d \mid \beta \in -(\mathbb{N}+1)\underline{a}_j + \operatorname{qdeg}(S_A/(t^{\underline{a}_j}))\}.$$

Consider as well the torus $\mathbb{G}_m^d := \operatorname{Spec}(\mathbb{C}[t_1^\pm, \dots, t_d^\pm])$, together with the torus embedding

$$\begin{aligned} h : \mathbb{G}_m^d &\to \mathbb{A}^N \\ (t_1, \dots, t_d) &\mapsto (t^{\underline{a}_1}, \dots, t^{\underline{a}_N}). \end{aligned}$$

The following proposition is a slight generalization of the results of Schulze and Walther [2009, Theorem 3.6, Corollary 3.8].

Proposition 4.3 [Reichelt and Sevenheck 2015, Proposition 3.11]. *Let A be a $d \times N$ integer matrix satisfying $\mathbb{Z}A = \mathbb{Z}^d$ and $\mathbb{N}A = \mathbb{Z}^d \cap \mathbb{R}_{\geq 0}A$. Assume that $\beta \notin \operatorname{sRes}(A)$. Then*

$$\mathcal{H}^0(h_+\mathcal{O}_{\mathbb{G}_m^d}^\beta) \cong \check{\mathcal{M}}_A^\beta,$$

where $\mathcal{O}_{\mathbb{G}_m^d}^\beta \cong \mathcal{D}_{\mathbb{G}_m^d}/\mathcal{D}_{\mathbb{G}_m^d} \cdot (\partial_{t_1}t_1 + \beta_1, \dots, \partial_{t_d}t_d + \beta_d)$.

For $\beta \in \mathbb{R}^d$, the $\mathcal{D}$-module $\mathcal{O}^{\beta}_{\mathbb{G}^d_m}$ underlies the complex mixed Hodge module ${}^p\mathbb{C}^{H,\beta}_{\mathbb{G}^d_m}$. Hence for $\beta \in \mathbb{R}^d \setminus \mathrm{sRes}(A)$ the $\mathcal{D}$-module $\check{\mathcal{M}}^{\beta}_A$ underlies the complex mixed Hodge module $\mathcal{H}^0 h_* {}^p\mathbb{C}^{H,\beta}_{\mathbb{G}^d_m}$. The Hodge filtration on $\check{\mathcal{M}}^{\beta}_A$ can be explicitly computed, provided that β belongs to a certain set $\mathfrak{A}_A$ of so-called admissible parameters β. We recall its definition from [Reichelt and Sevenheck 2015, Formula (14) right before Lemma 4.4]: let $\underline{c} := \underline{a}_1 + \cdots + \underline{a}_N$ and define for all facets F of $\mathbb{R}_{\geq 0}A$ the uniquely determined primitive, inward-pointing, normal vector $\underline{n}_F$ of F, such that $\langle \underline{n}_F, F\rangle = 0$ and $\langle \underline{n}_F, \mathbb{N}A\rangle \subset \mathbb{Z}_{\geq 0}$. Set $e_F := \langle \underline{n}_F, \underline{c}\rangle \in \mathbb{Z}_{>0}$. The set of admissible parameters of A is then defined by

$$\mathfrak{A}_A := \bigcap_{F \text{ facet}} \left\{\mathbb{R}\cdot F - \left[0, \tfrac{1}{e_F}\right) \cdot \underline{c}\right\}.$$

Theorem 4.4 [Reichelt and Sevenheck 2015, Theorem 4.17]. *For $\beta \in \mathfrak{A}_A$ the Hodge filtration on $\check{\mathcal{M}}^{\beta}_A$ is equal to the order filtration shifted by $N-d$, i.e.,*

$$F^H_{p+N-d}\check{\mathcal{M}}^{\beta}_A = F^{\mathrm{ord}}_p \check{\mathcal{M}}^{\beta}_A.$$

Let us define the cyclic $\mathcal{R}_{\mathbb{A}^1\times\mathbb{A}^N}$-module $\check{\mathcal{N}}^{\beta}_A := \mathcal{R}_{\mathbb{A}^1\times\mathbb{A}^N}/\mathcal{J}_z$, where $\mathcal{J}_z$ is the left ideal generated by

$$\check{E}^z_k = \sum_{i=1}^N a_{ki} z\partial_{w_i} w_i + z\beta_k, \quad \text{for } k = 1, \dots, d,$$

and

$$\check{\Box}_\ell = \prod_{\ell_i>0} w_i^{\ell_i} - \prod_{\ell_i<0} w_i^{-\ell_i}, \quad \text{for } l \in \mathbb{L}_A.$$

We will denote by $\check{M}^{\beta}_A := \Gamma(\mathbb{A}^N, \check{\mathcal{M}}^{\beta}_A)$ and $\check{N}^{\beta}_A := \Gamma(\mathbb{A}^1\times\mathbb{A}^N, \check{\mathcal{N}}^{\beta}_A)$ the modules of global sections of $\check{\mathcal{M}}^{\beta}_A$ and $\check{\mathcal{N}}^{\beta}_A$, respectively.

We will also consider the Rees module of $\check{M}^{\beta}_A$ with respect to the order filtration $F^{\mathrm{ord}}_\bullet$, which is given by $R^{F^{\mathrm{ord}}}\check{M}^{\beta}_A := \sum_{k\geq 0} z^k F^{\mathrm{ord}}_k \check{M}^{\beta}_A$. An easy computation shows $R^{F^{\mathrm{ord}}}\check{M}^{\beta}_A = \check{N}^{\beta}_A$, hence

$$R^{F^H}\check{M}^{\beta}_A = z^{N-d}\check{N}^{\beta}_A. \tag{3}$$

Definition 4.5. The $\mathcal{R}$-GKZ-hypergeometric system $\mathcal{N}^{\beta}_A$ is the cyclic $\mathcal{R}^{\mathrm{int}}_{\mathbb{A}^1\times\hat{\mathbb{A}}^N}$-module $\mathcal{R}^{\mathrm{int}}_{\mathbb{A}^1\times\hat{\mathbb{A}}^N}/\mathcal{I}$, where the left ideal $\mathcal{I}$ is generated by

$$E^z_0 := z^2\partial_z + \sum_{i=1}^N \lambda_i z\partial_{\lambda_i},$$

$$E^z_k := \sum_{i=1}^N a_{ki}\lambda_i z\partial_{\lambda_i} - z\beta_k, \quad \text{for } k = 1, \dots, d,$$

and

$$\Box^z_\ell := \prod_{\ell_i>0} (z\partial_{\lambda_i})^{\ell_i} - \prod_{\ell_i<0} (z\partial_{\lambda_i})^{-\ell_i}, \text{ for } \ell \in \mathbb{L}_A.$$

Remark 4.6. Note that, considering $\check{\mathcal{N}}_A^\beta$ as an $\mathcal{R}^{\mathrm{int}}_{\mathbb{A}^1\times\hat{\mathbb{A}}^N}$-module with the trivial action of $z^2\partial_z$, $\mathcal{N}_A^\beta$ is its Fourier–Laplace transform as $\mathcal{R}^{\mathrm{int}}_{\mathbb{A}^1\times\hat{\mathbb{A}}^N}$-modules, according to Remark 3.3.

Theorem 4.7. *Let A be a $d\times N$-matrix and $\beta\in\mathfrak{A}_A$ an admissible parameter. The $\mathcal{R}$-GKZ-hypergeometric system $z^{-d}\mathcal{N}_A^\beta$ underlies an algebraic, integrable, mixed twistor $\mathcal{D}$-module ${}^{\mathcal{T}}\mathcal{M}_A^\beta$.*

Proof. By the remark above, we know that $\mathcal{N}_A^\beta=\mathrm{FL}(\check{\mathcal{N}}_A^\beta)$, which in turn, thanks to the choice of β, Theorem 4.4 and formula (3), is equal to $\mathrm{FL}(z^{d-N}\mathcal{R}^{F^H}\check{\mathcal{M}}_A^\beta)$. Since $\mathcal{R}^{F^H}\check{\mathcal{M}}_A^\beta$ is the Rees module of a mixed Hodge module on $\mathbb{A}^N$, it gives rise to an algebraic, integrable mixed twistor $\mathcal{D}$-module on $\mathbb{A}^N$, say ${}^{\mathcal{T}}\check{\mathcal{M}}_A^\beta$. Then we can apply Proposition 3.5 and get

$$\mathcal{N}_A^\beta=z^{d-N}\,\mathrm{FL}(\mathrm{For}_2({}^{\mathcal{T}}\check{\mathcal{M}}_A^\beta))=z^d\,\mathrm{For}_2(\mathrm{FL}_{\mathrm{MTM}}({}^{\mathcal{T}}\check{\mathcal{M}}_A^\beta)).$$

The result follows from writing ${}^{\mathcal{T}}\mathcal{M}_A^\beta:=\mathrm{FL}_{\mathrm{MTM}}({}^{\mathcal{T}}\check{\mathcal{M}}_A^\beta)$. □

Corollary 4.8. *The analytification of ${}^{\mathcal{T}}\mathcal{M}_A^\beta$ gives rise to an irregular mixed Hodge module on $\mathbb{A}^N$ which has a natural extension to an $\mathcal{R}^{\mathrm{int}}_{\mathbb{A}^1\times\mathbb{P}^N}$-module underlying an object of* $\mathrm{IrrMHM}(\mathbb{P}^N)$.

Proof. This follows from applying [Sabbah 2018, Corollary 0.5] to the operations performed to get ${}^{\mathcal{T}}\mathcal{M}_A^\beta$. □

5. Application to confluent hypergeometric systems

In this section we are going to use the results achieved so far for the special case of the matrix

$$A=\left(\begin{array}{c|c|c}\underline{1}_m & \underline{0}_{m\times(n-1)} & \mathrm{Id}_m\\ \hline \underline{1}_{n-1} & -\mathrm{Id}_{n-1} & \underline{0}_{(n-1)\times m}\end{array}\right).$$

For the sake of simplicity, we will write $N=n+m$ in what follows. Before going on, let us introduce the main object of study of this section and state some of its basic properties, extending what we mentioned in the introduction.

Definition 5.1. Let $(n,m)\neq(0,0)$ be a pair of nonnegative integers, and let $\alpha_1,\dots,\alpha_n$ and $\beta_1,\dots,\beta_m$ be elements of $\mathbb{C}$. The hypergeometric $\mathcal{D}$-module of type (n,m) associated with the α_i and the β_j is defined as the quotient of $\mathcal{D}_{\mathbb{G}_m}$ by the left ideal generated by the so-called hypergeometric operator

$$\prod_{i=1}^n(t\partial_t-\alpha_i)-t\prod_{j=1}^m(t\partial_t-\beta_j).$$

We will denote it by $\mathcal{H}(\alpha_i;\beta_j)$.

Proposition 5.2. *Let $\mathcal{H}:=\mathcal{H}(\alpha_i;\beta_j)$ be a hypergeometric $\mathcal{D}$-module of type (n,m), and let η be any complex number. Then we have the following*:

(1) *If we denote the Kummer $\mathcal{D}$-module $\mathcal{D}_{\mathbb{G}_m}/(t\partial_t-\eta)$ by $\mathcal{K}_\eta$, then $\mathcal{H}\otimes_{\mathcal{O}_{\mathbb{G}_m}}\mathcal{K}_\eta\cong\mathcal{H}(\alpha_i+\eta;\beta_j+\eta)$. In particular, an overall integer shift of the parameters gives us an isomorphic $\mathcal{D}$-module.*

(2) *$\mathcal{H}$ is irreducible if and only if for any pair (i,j) of indices, $\alpha_i-\beta_j$ is not an integer.*

(3) *If $\mathcal{H}$ is irreducible, its isomorphism class depends only on the classes modulo $\mathbb{Z}$ of the α_i and the β_j, so we can choose such parameters on a fundamental domain of $\mathbb{C}/\mathbb{Z}$.*

Proof. A simple calculation shows (1). (2) follows from [Katz 1990, Propositions 2.11.9 and 3.2], whereas (3) is part of [ibid., Proposition 3.2]. □

As we mentioned in the introduction, we can express any one-dimensional hypergeometric $\mathcal{D}$-module as the inverse image of a GKZ hypergeometric $\mathcal{D}$-module (see [Castaño Domínguez and Sevenheck 2019, Corollary 2.9]). Notice that there is a similar statement at the level of $\mathcal{R}$-modules (see [ibid., Lemma 2.12]), yielding a description of the $\mathcal{R}^{\text{int}}_{\mathbb{A}^1_z\times\mathbb{G}_{m,t}}$-module $\widehat{\mathcal{H}}$ from Theorem 5.7 below as an inverse image of a GKZ-hypergeometric $\mathcal{R}$-module (as defined in [ibid., Definition 2.10]).

Proposition 5.3. *Let $\mathcal{H}(\alpha_i;\beta_j)$ be a hypergeometric $\mathcal{D}_{\mathbb{G}_m}$-module of type (n,m) with $\alpha_1=0$, let $A\in \mathrm{M}((N-1)\times N,\mathbb{Z})$ as above, and let $\gamma=(\beta_1,\dots,\beta_m,\alpha_2,\dots,\alpha_n)^{\mathrm{t}}$. Let $\iota:\mathbb{G}_m\to\mathbb{A}^N$ be given by $t\mapsto(t,1\dots,1)$. Then*

$$\mathcal{H}(\alpha_i;\beta_j)\cong\iota^+\mathcal{M}^\gamma_A.$$

Since the restriction map ι is not smooth we do not know a priori whether taking inverse image by it preserves irregular mixed Hodge modules. In order to show that $\mathcal{H}(\alpha_i;\beta_j)$ can be upgraded to an element of $\mathrm{IrrMHM}(\mathbb{G}_m)$ we use Proposition 3.8, where the reduction procedure is built in by the use of the Fourier kernel $\psi=w_1\cdot t+w_2+\cdots+w_N$.

Let $A\in\mathrm{M}((N-1)\times N,\mathbb{Z})$ as above and $\gamma=(\gamma_1,\dots,\gamma_{N-1})^{\mathrm{t}}\in\mathfrak{A}_A$. The $\mathcal{D}_{\mathbb{A}^N}$-module $\check{\mathcal{M}}^\gamma_A$ underlies a mixed Hodge module on $\mathbb{A}^N$, so that the Rees module $\mathcal{R}^{F^H}(\check{\mathcal{M}}^\gamma_A)$ then gives rise to an algebraic, integrable mixed twistor $\mathcal{D}$-module on $\mathbb{A}^N$ that we denote by ${}^{\mathcal{T}}\check{\mathcal{M}}^\gamma_A$. Then we have the following concrete description of its Fourier–Laplace transform $\mathrm{FL}^\psi_{\mathrm{MTM}}({}^{\mathcal{T}}\check{\mathcal{M}}^\gamma_A)=q_*(p^*({}^{\mathcal{T}}\check{\mathcal{M}}^\gamma_A)\otimes\mathcal{T}^{\psi/z})$.

Proposition 5.4. *Let A and γ be as before. Then the $\mathcal{R}^{\text{int}}_{\mathbb{A}^1\times\mathbb{G}_m}$-module $\mathrm{For}_2(\mathrm{FL}^\psi_{\mathrm{MTM}}({}^{\mathcal{T}}\check{\mathcal{M}}^\gamma_A))$ can be expressed as $\mathcal{R}^{\text{int}}_{\mathbb{A}^1\times\mathbb{G}_m}/(P,H)$, where*

$$P=z^2\partial_z+(n-m)tz\partial_t+\varepsilon z\quad\textit{and}\quad H=zt\partial_t\prod_{i=1}^{n-1}z(t\partial_t-\gamma_{m+i})-t\prod_{j=1}^{m}z(t\partial_t-\gamma_j),$$

with $\varepsilon=\sum_{j=1}^m\gamma_j-\sum_{i=m+1}^{N-1}\gamma_i+N-1$.

Proof. As said after Theorem 4.4, for any γ inside the domain $\mathfrak{A}_A$ of admissible parameters, the Hodge filtration of $\check{\mathcal{M}}^\gamma_A$ is the order filtration shifted by $N-(N-1)=1$. Therefore, for such values of γ we can give an explicit expression of the Rees module of the filtered module $(\check{\mathcal{M}}^\gamma_A,F^H_\bullet)$. Namely, we have the isomorphism of $\mathcal{R}^{\text{int}}_{\mathbb{A}^1\times\mathbb{A}^N}$-modules

$$\mathcal{R}^{F^H}(\check{\mathcal{M}}^\gamma_A)\cong z\check{\mathcal{N}}^\gamma_A:=\mathcal{R}^{\text{int}}_{\mathbb{A}^1\times\mathbb{A}^N}/(\check{E}^z_i,\check{E}^z_j,\check{\Box},z^2\partial_z-z),$$

where

$$\check{E}_i^z = z\partial_{w_1} w_1 - z\partial_{w_i} w_i + \gamma_{m+i-1} z, \quad \text{for } i = 2, \dots, n,$$
$$\check{E}_j^z = z\partial_{w_1} w_1 + z\partial_{w_{n+j}} w_{n+j} + \gamma_j z, \quad \text{for } j = 1, \dots, m,$$
$$\check{\Box} = \prod_{i=1}^{n} w_i - \prod_{j=1}^{m} w_{n+j}.$$

First we compute $\mathrm{FL}^{\psi}(z\check{\mathcal{N}}_A^{\gamma})$, which involves performing three operations with $z\check{\mathcal{N}}_A^{\gamma}$: inverse image by $p : \mathbb{G}_m \times \mathbb{A}^N \to \mathbb{A}^N$, tensor product with the $\mathcal{R}^{\mathrm{int}}_{\mathbb{A}^1 \times \mathbb{G}_m \times \mathbb{A}^N}$-module $\mathcal{A}_{\mathrm{aff}}^{\psi/z}$ and direct image by $q : \mathbb{G}_m \times \mathbb{A}^N \to \mathbb{G}_m$. The first one is pretty easy. Namely

$$p^+ z\check{\mathcal{N}}_A^{\gamma} \cong \mathcal{R}^{\mathrm{int}}_{\mathbb{A}^1 \times \mathbb{G}_m \times \mathbb{A}^N} / (\check{E}_i^z, \check{E}_j^z, \check{\Box}, z^2\partial_z - z, z\partial_t).$$

Let us tensor now $p^+ z\check{\mathcal{N}}_A^{\gamma}$ with $\mathcal{A}_{\mathrm{aff}}^{\psi/z}$. This $\mathcal{R}^{\mathrm{int}}$-module can be presented as $\mathcal{R}^{\mathrm{int}}_{\mathbb{A}^1 \times \mathbb{G}_m \times \mathbb{A}^N} \cdot e^{\psi/z} = \mathcal{R}^{\mathrm{int}}_{\mathbb{A}^1 \times \mathbb{G}_m \times \mathbb{A}^N} / \mathcal{I}^{\psi}$, where $\mathcal{I}^{\psi}$ is the left ideal generated by

$$z^2\partial_z + w_1 t + w_2 + \cdots + w_N, \quad z\partial_t - w_1, \quad z\partial_{w_1} - t, \quad z\partial_{w_i} - 1, i = 2, \dots, N.$$

For $n \in p^+ z\check{\mathcal{N}}_A^{\gamma}$, we will call n^{ψ} the tensor $n \otimes e^{\psi/z}$. Then we can obtain the formulas

$$(z\partial_{w_1} w_1 n \otimes e^{\psi/z}) = z\partial_{w_1}(w_1 n \otimes e^{\psi/z}) - t(n \otimes w_1 e^{\psi/z}) = (z\partial_{w_1} w_1 - tz\partial_t) \cdot n^{\psi},$$
$$(z\partial_{w_k} w_k n \otimes e^{\psi/z}) = z\partial_{w_k}(w_k n \otimes e^{\psi/z}) - (n \otimes w_k e^{\psi/z}) = (z\partial_{w_k} w_k - w_k) \cdot n^{\psi}, \quad \text{for } k = 2, \dots, N,$$
$$(z^2\partial_z n \otimes e^{\psi/z}) = z^2\partial_z \cdot n^{\psi} - (n \otimes (-\psi) e^{\psi/z}) = (z^2\partial_z + w_1 t + w_2 + \cdots + w_N) \cdot n^{\psi},$$
$$(z\partial_t n \otimes e^{\psi/z}) = z\partial_t \cdot n^{\psi} - (n \otimes w_1 e^{\psi/z}) = (z\partial_t - w_1) \cdot n^{\psi}.$$

Hence $p^+ z\check{\mathcal{N}}_A^{\gamma} \otimes \mathcal{A}_{\mathrm{aff}}^{\psi/z}$ is the cyclic $\mathcal{R}^{\mathrm{int}}_{\mathbb{A}^1 \times \mathbb{G}_m \times \mathbb{A}^N}$-module $\mathcal{R}^{\mathrm{int}}_{\mathbb{A}^1 \times \mathbb{G}_m \times \mathbb{A}^N} / \mathcal{J}^{\psi}$, with $\mathcal{J}^{\psi}$ being the left ideal generated by

$$\prod_{i=1}^{n} w_i - \prod_{j=1}^{m} w_{n+j}, \quad z^2\partial_z - z + w_1 t + w_2 + \cdots + w_N, \quad z\partial_t - w_1,$$
$$z\partial_{w_1} w_1 - tz\partial_t - z\partial_{w_i} w_i + w_i + \gamma_{m+i-1} z, \quad \text{for } i = 2, \dots, n,$$
$$z\partial_{w_1} w_1 - tz\partial_t + z\partial_{w_{n+j}} w_{n+j} - w_{n+j} + \gamma_j z, \quad \text{for } j = 1, \dots, m.$$

We now consider the zeroth cohomology $\mathcal{H}^0 q_+(p^+ z\check{\mathcal{N}}_A^{\gamma} \otimes \mathcal{A}_{\mathrm{aff}}^{\psi/z})$, which is in turn the N-th cohomology of the de Rham complex $q_* \mathrm{DR}_{\mathbb{A}^1 \times \mathbb{G}_m \times \mathbb{A}^N / \mathbb{A}^1 \times \mathbb{G}_m}(p^+ z\check{\mathcal{N}}_A^{\gamma} \otimes \mathcal{A}_{\mathrm{aff}}^{\psi/z})$. This is given by the cyclic $\mathcal{R}^{\mathrm{int}}_{\mathbb{A}^1 \times \mathbb{G}_m}$-module $\mathcal{R}^{\mathrm{int}}_{\mathbb{A}^1 \times \mathbb{G}_m} / (P', H')$, where the operators P' and H' are given by

$$P' := z^2\partial_z + (n-m) tz\partial_t + \varepsilon' z, \quad H' := zt\partial_t \prod_{i=1}^{n-1} (zt\partial_t - \gamma_{m+i} z) - (-1)^m t \prod_{j=1}^{m} (zt\partial_t - \gamma_j z)$$

and $\varepsilon' := \sum_{j=1}^{m} \gamma_j - \sum_{i=m+1}^{N-1} \gamma_i - 1$. Replacing t by $(-1)^m t$ we obtain that $\mathrm{FL}^{\psi}(z\check{\mathcal{N}}_A^{\gamma}) \cong \mathcal{R}^{\mathrm{int}}_{\mathbb{A}^1\times\mathbb{G}_m}/(P', H)$, with

$$H := zt\partial_t \prod_{i=1}^{n-1}(zt\partial_t - \gamma_{m+i}z) - t\prod_{j=1}^{m}(zt\partial_t - \gamma_j z).$$

Now it follows from Proposition 3.8 that

$$\mathrm{For}_2(\mathrm{FL}^{\psi}_{\mathrm{MTM}}({}^{\mathcal{T}}\check{\mathcal{M}}_A^{\gamma})) \cong z^{-N}\,\mathrm{FL}^{\psi}(z\check{\mathcal{N}}_A^{\gamma}) \cong \mathcal{R}^{\mathrm{int}}_{\mathbb{A}^1\times\mathbb{G}_m}/(P, H)$$

with

$$P = z^2\partial_z + (n-m)tz\partial_t + \varepsilon z \quad \text{and} \quad H = zt\partial_t \prod_{i=1}^{n-1} z(t\partial_t - \gamma_{m+i}) - t\prod_{j=1}^{m} z(t\partial_t - \gamma_j),$$

and $\varepsilon = \sum_{j=1}^{m} \gamma_j - \sum_{i=m+1}^{N-1} \gamma_i + N - 1$. □

Remark 5.5. As a matter of fact, we do not have to restrict ourselves to the region $\mathfrak{A}_A$ to find our admissible parameters. If we have $\gamma \in \mathfrak{A}_A$ and add to it an integer vector $\underline{k} \in \mathbb{Z}^{N-1}$ with no negative entries, then $\gamma + \underline{k} \notin \mathrm{sRes}(A)$ by definition (see the proof of [Reichelt and Sevenheck 2015, Lemma 4.5]). Therefore, since $\mathcal{O}^{\gamma}_{\mathbb{G}_m^d} \cong \mathcal{O}^{\gamma+\underline{k}}_{\mathbb{G}_m^d}$ for any integer vector $\underline{k}$, we have $\check{\mathcal{M}}_A^{\gamma} \cong \check{\mathcal{M}}_A^{\gamma+\underline{k}}$ by Proposition 4.3 and the statement of the proposition holds true after changing $\mathfrak{A}_A$ by $\mathfrak{A}_A + \mathbb{N}^{N-1}$.

We will also make use of the following result, which calculates the admissible domain $\mathfrak{A}_A$ for the matrix A in our particular context.

Lemma 5.6. *Let $A \in M((N-1)\times N, \mathbb{Z})$ be the matrix defined at the beginning of the section. Consider a point $p = (p_1, \dots, p_m, q_1, \dots, q_{n-1}) \in [0, 1)^{N-1}$. Let us define*

$$p_- := \min((\{p_1, \dots, p_m\} \setminus \{0\}) \cup \{1\}) \quad \textit{and} \quad p_+ := \max\{p_1, \dots, p_m\},$$

that is, the minimum of the p_i that do not vanish (taking $p_- = 1$ if all of them are zero) and the maximum of them all.

Then, p belongs to $(\mathfrak{A}_A + \mathbb{N}^{N-1}) \subset \mathbb{R}^{N-1}$ if and only if, for all $i = 1, \dots, n-1$

- *$q_i \in [0, p_-)$ if some p_i vanishes, or*
- *$q_i \in [0, p_-) \cup [p_+, 1)$, otherwise.*

Proof. We will first find the expression for the admissible region $\mathfrak{A}_A$. For this purpose, we must find a set of hyperplanes containing the facets of the cone $C := \mathbb{R}_{\geq 0}A \subset \mathbb{R}^{N-1}$. Denote by $\{\underline{u}_1, \dots, \underline{u}_{N-1}\}$ the canonical basis of $\mathbb{R}^{N-1}$ and write $x_1, \dots, x_{N-1}$ for the corresponding coordinates.

Since any face of a cone is generated by a subset of its generators, and for our given matrix A, any $(N-1)\times(N-1)$-minor is nonzero (so that any subset of $N-1$ columns generates a full-dimensional cone), we see that any facet can contain at most $N-2$ columns. On the other hand, such facet must be $(N-2)$-dimensional, so it cannot be generated by fewer columns. Therefore, we can conclude that it contains exactly $N-2$ columns.

Any linear functional h defining a facet of C must satisfy that $h(C) \geq 0$. Denote by $H_{k,l}$ the hyperplane not containing $\underline{a}_k$ and $\underline{a}_\ell$. There are five classes of these hyperplanes: $H_{1,i}$, $H_{1,n+j}$, H_{i_1,i_2}, $H_{i,n+j}$, $H_{n+j_1,n+j_2}$ with $i, i_1, i_2 \in \{2, \ldots, n\}$ and $j, j_1, j_2 \in \{1, \ldots, m\}$. The linear functionals defining them are, respectively,

$$\begin{aligned} h_{1,i} &:= x_{m+i-1}, \\ h_{1,n+j} &:= x_j, \\ h_{i_1,i_2} &:= x_{m+i_1-1} - x_{m+i_2-1}, \\ h_{i,n+j} &:= x_j - x_{m+i-1}, \\ h_{n+j_1,n+j_2} &:= x_{j_1} - x_{j_2}. \end{aligned}$$

All of the linear forms $h_{1,i}$, h_{i_1,i_2} and $h_{n+j_1,n+j_2}$ (for the corresponding values of i, i_1, i_2, j_1, j_2) take both negative and positive values on some columns of A, so the associated hyperplanes do not contain any facet.

We conclude that each facet of C is contained in one of the following hyperplanes:

$$\begin{aligned} H_{1,n+j} : x_j = 0 \quad & \text{for } j = 1, \ldots, m, \\ H_{i,n+j} : x_j - x_{m+i-1} = 0 \quad & \text{for } i = 2, \ldots, n,\ j = 1, \ldots, m. \end{aligned} \tag{4}$$

These hyperplanes are different from each other and the respective functionals satisfy $h_{1,n+j}(C) \geq 0$ and $h_{i,n+j}(C) \geq 0$. Hence each of them contains a different facet of the cone C.

The primitive, inward-pointing normal vectors of the hyperplanes $H_{1,n+j}$ and $H_{i,n+j}$ are $\underline{n}_{1,n+j} := \underline{u}_j$ and $\underline{n}_{i,n+j} := \underline{u}_j - \underline{u}_{m+i-1}$, respectively. Denote by $\underline{c}$ the sum of all columns of A. We have $\underline{c} = 2(\underline{u}_1 + \cdots + \underline{u}_m)$ and $e_{k,l} := \langle \underline{n}_{k,l}, \underline{c} \rangle = 2$, where k and l take the admissible values corresponding to the hyperplanes we consider in (4) (i.e., we have either $(k,l) = (1, n+j)$ or $(k,l) = (i, n+j)$ for $i = 2, \ldots, n$ and $j = 1, \ldots, m$). Define

$$\begin{aligned} \mathfrak{A}_{k,l} &:= H_{k,l} - \left[0, \tfrac{1}{e_{k,l}}\right) \cdot \underline{c} \\ &= H_{k,l} - [0, 1) \cdot (\underline{u}_1 + \cdots + \underline{u}_m) \\ &= \begin{cases} H_{1,n+j} - [0,1) \cdot u_j & \text{for } j = 1, \ldots, m, \\ H_{i,n+j} - [0,1) \cdot u_j & \text{for } i = 2, \ldots, n,\ j = 1, \ldots, m, \end{cases} \end{aligned}$$

since for $(k,l) = (1, n+j)$ and $(k,l) = (i, n+j)$, the vectors $\underline{u}_1, \ldots, \underline{u}_{j-1}, \underline{u}_{j+1}, \ldots, \underline{u}_m$ are contained in $H_{1,n+j}$ and $H_{i,n+j}$, respectively. Then we have

$$\mathfrak{A}_{1,n+j} = H_{1,n+j} - [0,1) \cdot u_j = \{(x_1, \ldots, x_{N-1}) \in \mathbb{R}^{N-1} \mid -1 < x_j \leq 0\}$$

for all $j = 1, \ldots, m$ and

$$\mathfrak{A}_{i,n+j} = H_{i,n+j} - [0,1) \cdot u_j = \{(x_1, \ldots, x_{N-1}) \in \mathbb{R}^{N-1} \mid -1 < x_j - x_{m+i-1} \leq 0\}$$

for all $i = 2, \ldots, n$, $j = 1, \ldots, m$. According to the construction given before Theorem 4.4, we can conclude that

$$\mathfrak{A}_A = \bigcap_{F \text{ facet}} \{\mathbb{R} \cdot F - [0, \tfrac{1}{e_F}) \cdot \underline{c}\} = \bigcap_{k,l \text{ from (4)}} \mathfrak{A}_{k,l},$$

so we can describe the admissible region $\mathfrak{A}_A$ as

$$\mathfrak{A}_A : \begin{cases} -1 < x_j \leq 0 & \text{for } j = 1, \ldots, m, \\ -1 < x_j - x_{m+i-1} \leq 0 & \text{for } i = 2, \ldots, n,\ j = 1, \ldots, m \end{cases} \subset \mathbb{R}^{N-1}.$$

Now let us pick a point $p \in [0, 1)^{N-1} \cap (\mathfrak{A}_A + \mathbb{N}^{N-1})$, and take $\underline{k} = (k_1, \ldots, k_{N-1}) \in \mathbb{N}^{N-1}$ such that $p \in [0, 1)^{N-1} \cap (\mathfrak{A}_A + \underline{k})$. The shifted domain is given by

$$\mathfrak{A}_A + \underline{k} : \begin{cases} -1 + k_j < x_j \leq k_j & \text{for } j = 1, \ldots, m, \\ -1 + k_j - k_{m+i-1} < x_j - x_{m+i-1} \leq k_j - k_{m+i-1} & \text{for } i = 2, \ldots, n,\ j = 1, \ldots, m \end{cases} \subset \mathbb{R}^{N-1}.$$

Assume first there is a vanishing coordinate p_{j_0}. Then we must have $k_{j_0} = 0$. For such an index and any $i = 1, \ldots, n-1$, we can consider the $n-1$ inequalities

$$-1 - k_{m+i} < -q_i \leq -k_{m+i},$$

from where we deduce that every q_i belongs to $[k_{m+i}, k_{m+i} + 1) \cap [0, 1)$, for $i = 1, \ldots, n-1$. In order for those intersections to be nonempty, we must have $k_{m+i} + 1 > 0$ and $k_{m+i} < 1$, so necessarily $k_{m+i} = 0$ for all i (and hence q_i must lie within $[0, 1)$, which is no new information).

Now, for any nonvanishing p_j, it is clear that $k_j = 1$. Then, if we look at the remaining inequalities, we see that

$$0 < p_j - q_i \leq 1,$$

for every $i = 1, \ldots, n-1$, and any $j \in \{1, \ldots, m\}$ such that $p_j \neq 0$. Therefore, every q_i belongs to $[0, 1) \cap \bigcap_{p_j \neq 0} [p_j - 1, p_j) = [0, p_-)$. Obviously, if $p_j = 0$ for all $j = 1, \ldots, m$, we obtain that the q_i belong all to $[0, 1) = [0, p_-)$.

Assume now that no p_j vanishes. Then $k_1 = \ldots = k_m = 1$. It follows that we can express the shifted region $\mathfrak{A}_A + \underline{k}$ as

$$\mathfrak{A}_A + \underline{k} : \begin{cases} 0 < x_j \leq 1 & \text{for } j = 1, \ldots, m, \\ -k_{m+i-1} < x_j - x_{m+i-1} \leq 1 - k_{m+i-1} & \text{for } i = 2, \ldots, n,\ j = 1, \ldots, m \end{cases} \subset \mathbb{R}^{N-1}.$$

Then, for any $j = 1, \ldots, m$, we have $q_i \in [0, 1) \cap [p_j + k_{m+i} - 1, p_j + k_{m+i})$, for $i = 1, \ldots, n-1$. As before, this implies that $p_j + k_{m+i} > 0$ and $p_j + k_{m+i} - 1 < 1$, for each $j = 1, \ldots, m$. Since each p_j lives in $(0, 1)$, the k_{m+i-1} can only be either 0 or 1.

Pick an $i \in \{1, \ldots, n-1\}$ such that $k_{m+i} = 0$. Then, as before,

$$q_i \in \bigcap_{j=1}^{m} [p_j - 1, p_j) \cap [0, 1) = [0, p_-).$$

If our index i is such that $k_{m+i} = 1$, then

$$q_i \in \bigcap_{j=1}^{m} [p_j, p_j + 1) \cap [0, 1) = [p_+, 1),$$

and one direction of the statement is done.

To show the other implication of the lemma, suppose now that every q_i lies within $[0, p_-) \cup [p_+, 1)$ for $i = 1, \dots, n-1$, and no p_j vanishes. We can rewrite this as a disjunction: either $q_i \in \bigcap_{j=1}^{m} [0, p_j) = [0, 1) \cap \bigcap_{j=1}^{m} [p_j - 1, p_j)$ or $q_i \in \bigcap_{j=1}^{m} [p_j, 1) = [0, 1) \cap \bigcap_{j=1}^{m} [p_j, p_j + 1)$. If $q_i \in [0, p_-)$, define $k_{m+i} := 0$. Otherwise, we take $k_{m+i} := 1$. Summing up, it is clear that

$$p \in (\mathfrak{A}_A + (1, \overset{(m)}{\dots}, 1, k_{m+1}, \dots, k_{N-1})) \cap [0, 1)^{N-1}.$$

If some p_j vanishes, and every q_i belongs to $[0, p_-)$, we can do the same as above to see that

$$p \in (\mathfrak{A}_A + (k_1, \dots, k_m, 0, \dots, 0)) \cap [0, 1)^{N-1},$$

where k_j vanishes if so does p_j and is equal to 1 if $p_j \neq 0$. □

As a consequence of the above calculation of the set of admissible parameters, let us prove a result extending [Castaño Domínguez and Sevenheck 2019, Theorem 2.13].

Theorem 5.7. *Let $\alpha_1, \dots, \alpha_n$ and $\beta_1, \dots, \beta_m$ be real numbers, lying on the interval $[0, 1)$ and increasingly ordered. Assume moreover that*:

- *No difference $\alpha_i - \beta_j$ is zero, for any $i = 1, \dots, n$ and $j = 1, \dots, m$.*
- *After applying the bijection $[0, 1) \to S^1$ given by $x \mapsto e^{2\pi i x}$, all the images of the α_i are at one arc of the unit circle, while those of the β_j find themselves at the complementary arc. (In other words and going back to the interval $[0, 1)$, either no α_i belongs to any interval (β_j, β_{j+1}) or vice versa.)*

Consider the operators P and H given by

$$P = z^2 \partial_z + (n-m) t z \partial_t + \varepsilon z \quad \textit{and} \quad H = \prod_{i=1}^{n} z(t\partial_t - \alpha_i) - t \prod_{j=1}^{m} z(t\partial_t - \beta_j),$$

with $\varepsilon = \sum_{j=1}^{m} \beta_j - \sum_{i=1}^{n} \alpha_i + N - 1$. Let $\widehat{\mathcal{H}}(\alpha_i; \beta_j)$ be the $\mathcal{R}^{\mathrm{int}}_{\mathbb{A}^1_z \times \mathbb{G}_m}$-module

$$\widehat{\mathcal{H}}(\alpha_i; \beta_j) := \mathcal{O}_{\mathbb{A}^1_z \times \mathbb{G}_m} \langle z^2 \partial_z, z t \partial_t \rangle / (P, H).$$

Then, $\widehat{\mathcal{H}}(\alpha_i; \beta_j)$ underlies a unique object of $\mathrm{IrrMHM}(\mathbb{G}_m)$ *with associated $\mathcal{D}_{\mathbb{G}_m}$-module $\mathcal{H}(\alpha_i; \beta_j)$. It can be uniquely extended to an irreducible $\mathcal{R}^{\mathrm{int}}_{\mathbb{A}^1_z \times \mathbb{P}^1}$-module underlying an object of* $\mathrm{IrrMHM}(\mathbb{P}^1)$.

Proof. Let us assume first that $\alpha_1 = 0$. Then, by the first assumption on the α_i and the β_j, we have $\beta_j \neq 0$ for every j. By the second assumption we can deduce that no α_i is between any two β_j, but all of the β_j must be between two certain α_i. Thanks to Lemma 5.6, this means that $\gamma := (\beta_1, \dots, \beta_m, \alpha_2, \dots, \alpha_n)$

belongs to $\mathfrak{A}_A + \mathbb{N}^{N-1}$, where A is the matrix of the beginning of the section. As a consequence, by Proposition 5.4 and Remark 5.5 we have that

$$\mathrm{For}_2(\mathrm{FL}^{\psi}_{\mathrm{MTM}}({}^{\mathcal{T}}\check{\mathcal{M}}^{\gamma}_A)) \cong \widehat{\mathcal{H}}(\alpha_i; \beta_j)$$

(recall that ${}^{\mathcal{T}}\check{\mathcal{M}}^{\gamma}_A$ is the algebraic integrable mixed twistor $\mathcal{D}$-module with underlying $\mathcal{R}^{\mathrm{int}}_{\mathbb{A}^N}$-module $\mathcal{R}^{F^H}\check{\mathcal{M}}^{\gamma}_A$, i.e., such that $\mathrm{For}_2(\check{\mathcal{M}}^{\gamma}_A) = \mathcal{R}^{F^H}\check{\mathcal{M}}^{\gamma}_A)$. We have moreover that ${}^{\mathcal{T}}\check{\mathcal{M}}^{\gamma}_A \in \mathrm{IrrMHM}(\mathbb{A}^N)$ and thanks to [Sabbah 2018, Corollary 0.5], we know that the functors entering in the definition of $\mathrm{FL}^{\psi}_{\mathrm{MTM}}$ preserve the category of irregular mixed Hodge modules, so we conclude that $\widehat{\mathcal{H}}(\alpha_i; \beta_j)$ underlies an element of $\mathrm{IrrMHM}(\mathbb{G}_m)$.

Assume now that $\alpha_1 > 0$. For any real number η, denote by $\widehat{\mathcal{K}}_\eta$ the Kummer $\mathcal{R}_{\mathbb{A}^1\times\mathbb{G}_m}$-module $\mathcal{R}^{\mathrm{int}}_{\mathbb{A}^1\times\mathbb{G}_m}/(z^2\partial_z, tz\partial_t - z\eta)$.

The tensor product of $\mathcal{R}^{\mathrm{int}}_{\mathbb{A}^1\times\mathbb{G}_m}$-modules $\widehat{\mathcal{H}}(\alpha_i; \beta_j) \otimes_{\mathcal{O}_{\mathbb{A}^1\times\mathbb{G}_m}} \widehat{\mathcal{K}}_{-\alpha_1}$ gives rise to the corresponding tensor product of twistor $\mathcal{D}$-modules on $\mathbb{G}_m$. This product can be presented as $\widehat{\mathcal{H}}(\alpha'_i; \beta'_j)$, where $\alpha'_i = \alpha_i - \alpha_1$ for every i and $\beta'_j = \beta_j - \alpha_1$ for every j. The assumptions on the parameters imply that $\alpha'_1 = 0$ and the vector $(\beta'_1, \ldots, \beta'_m, \alpha'_2, \ldots, \alpha'_n)$ lives in $\mathfrak{A}_A + \mathbb{N}^{N-1}$. Then, arguing as before, such tensor product is an irregular mixed Hodge module of exponential-Hodge origin. Since $\widehat{\mathcal{K}}_{\alpha_1}$ is the faithful image of a mixed Hodge module on $\mathbb{G}_m$, the tensor product with it preserves the condition of being in $\mathrm{IrrMHM}(\mathbb{G}_m)$ due to [Sabbah 2018, Corollary 0.5], and so is the case of our original $\mathcal{R}^{\mathrm{int}}_{\mathbb{A}^1_z\times\mathbb{G}_m}$-module

$$\widehat{\mathcal{H}}(\alpha_i; \beta_j) \cong \widehat{\mathcal{H}}(\alpha'_i; \beta'_j) \otimes_{\mathcal{O}_{\mathbb{A}^1\times\mathbb{G}_m}} \widehat{\mathcal{K}}_{\alpha_1}.$$

This ends the statement on the existence. Let us prove now the claims on the unicity, as in [Castaño Domínguez and Sevenheck 2019, Theorem 2.13], noting that the condition on the differences $\alpha_i - \beta_j$ is equivalent to $\mathcal{H}$ being irreducible, and thus rigid (see [ibid., Proposition 2.5], noting that all the parameters belong to $[0, 1)$).

Consider now any twistor $\mathcal{D}$-module $\widehat{\mathcal{H}}'$ on $\mathbb{G}_{m,t}$ whose underlying $\mathcal{D}_{\mathbb{G}_{m,t}}$-module is $\mathcal{H}$. Since the functor Ξ_{DR} is faithful by [Mochizuki 2015a, Remark 7.2.9], we have an injection of Hom groups

$$\mathrm{Hom}_{\mathrm{MTM}(\mathbb{G}_{m,t})}(\widehat{\mathcal{H}}, \widehat{\mathcal{H}}') \hookrightarrow \mathrm{Hom}_{\mathcal{D}_{\mathbb{G}_{m,t}}}(\mathcal{H}, \mathcal{H}).$$

But $\mathcal{H}$ is irreducible, so its only endomorphism is the identity and then the twistor $\mathcal{D}$-module underlying $\mathcal{H}$ is unique.

On the other hand, let $j : \mathbb{G}_{m,t} \hookrightarrow \mathbb{P}^1$ be the canonical inclusion and consider the $\mathcal{D}_{\mathbb{P}^1}$-module $\mathcal{H}_{pr} := j_{\dagger+}\mathcal{H}$. It is an irreducible holonomic $\mathcal{D}_{\mathbb{P}^1}$-module, because so is $\mathcal{H}$ by the assumption on the α_i and the β_j. Then it gives rise to a unique pure integrable twistor $\mathcal{D}$-module $\widehat{\mathcal{H}}_{pr}$ on $\mathbb{P}^1$ by [Mochizuki 2011, Theorem 1.4.4; Sabbah 2018, Remark 1.39]. In addition, its underlying $\mathcal{D}_{\mathbb{P}^1}$-module $\mathcal{H}_{pr}$ is rigid, as $\mathcal{H}$ was. As a consequence, we can invoke [ibid., Theorem 0.7] and claim that such twistor $\mathcal{D}$-module on $\mathbb{P}^1$ is in fact an object of $\mathrm{IrrMHM}(\mathbb{P}^1)$. Take now $\widehat{\mathcal{H}}' := j^+\widehat{\mathcal{H}}_{pr}$, which is an irregular mixed Hodge module whose underlying $\mathcal{D}_{\mathbb{G}_{m,t}}$-module is $\mathcal{H}$, by [Mochizuki 2015a, Proposition 14.1.24]. Then we must have, as was just shown, $\widehat{\mathcal{H}}' \cong \widehat{\mathcal{H}}$, so that the extension $\widehat{\mathcal{H}}_{pr}$ of $\widehat{\mathcal{H}}$ is unique, and we are done. □

Remark 5.8. Let us consider the last theorem for the case $m=n$, that is, the case of regular hypergeometric systems. Consider $\widehat{\mathcal{H}}$ as a $\mathcal{R}_{\mathbb{A}^1_z\times\mathbb{G}_m}$-module only, as such it is isomorphic to $\mathcal{R}_{\mathbb{A}^1_z\times\mathbb{G}_m}/(H)$, where now

$$H=\prod_{i=1}^{m} z(t\partial_t-\alpha_i)-t\prod_{j=1}^{m} z(t\partial_t-\beta_j).$$

$\mathcal{R}_{\mathbb{A}^1_z\times\mathbb{G}_m}$ is graded by degree in z (where z has degree 1), and since H is homogenous (which is not the case if $n\neq m$), we see that $\widehat{\mathcal{H}}$ is a *graded* $\mathcal{R}_{\mathbb{A}^1_z\times\mathbb{G}_m}$-module. It is obviously *strict*, i.e., it has no z-torsion, and then by [Sabbah and Schnell 2018, A.7(5)], we see that $\widehat{\mathcal{H}}$ is the Rees module of a filtered $\mathcal{D}_{\mathbb{G}_m}$-module, namely, the (regular) hypergeometric module $\mathcal{H}(\alpha_i;\beta_j)$ together with the filtration by order of differential operators. Notice also that if $n=m$, we have $P=z^2\partial_z+\varepsilon z$, which implies that $\widehat{\mathcal{H}}$ has an action by $z\partial_z$ and that if we write $\widehat{\mathcal{H}}=\oplus_k\widehat{\mathcal{H}}_k$ (grading with respect to z), then for any $m\in\widehat{\mathcal{H}}_k$, we have $(z\partial_z)(m)=(k-\epsilon)m$.

Now suppose that we have $n=m$ and that additionally the hypotheses of the last theorem are satisfied, then since $\widehat{\mathcal{H}}(\alpha_i;\beta_j)$ is the unique object in IrrMHM($\mathbb{G}_m$) (lying actually in the essential image of MHM($\mathbb{G}_m$)) with underlying $\mathcal{D}_{\mathbb{G}_m}$-module $\mathcal{H}(\alpha_i;\beta_j)$, it is the Rees module of the filtered module $(\mathcal{H}(\alpha_i;\beta_j),F^H_\bullet)$, where $F^H_\bullet$ denotes the Hodge filtration of the complex variation of Hodge structures on $\mathcal{H}(\alpha_i;\beta_j)$. Hence $F^H_\bullet\mathcal{H}(\alpha_i;\beta_j)=F^{ord}_\bullet\mathcal{H}(\alpha_i;\beta_j)$ in this case. Moreover, if we put

$$R_k:=\prod_{i=1}^{k}(t\partial_t-\alpha_i)$$

for $k=0,\dots,n-1$ (where $R_0:=1$), then $(R_k)_{k=0,\dots,n-1}$ is an $\mathcal{O}_{\mathbb{G}_m}$-basis of $\mathcal{H}(\alpha_i,\beta_j)$ and yields a splitting of the Hodge filtration $F^H_\bullet$. In particular, we obtain that the Hodge numbers $h^p(\mathcal{H}(\alpha_i;\beta_j))=\dim(F^H_k/F^H_{k-1})$ are all equal to one. This is consistent with [Fedorov 2018, Theorem 1] (up to an overall shift, as noticed in that theorem) in the version of [Castaño Domínguez and Sevenheck 2019, Proposition 2.6], since under the assumption of Theorem 5.7, the function $\#\{j:\beta_j<\alpha_k\}$ is constant.

We will finish this section with a calculation of an irregular Hodge filtration, similar to the last section of [Castaño Domínguez and Sevenheck 2019]. In that reference, the authors computed such a filtration in the case where the hypergeometric $\mathcal{D}$-module had a purely irregular singularity at infinity, that is, it was of type $(n,0)$. It is immediate to see that for modules of type $(n,1)$, the second assumption of Theorem 5.7 holds true, so that we obtain an explicit description of the $\mathcal{R}^{\text{int}}_{\mathbb{A}^1_z\times\mathbb{G}_m}$-module underlying the irregular Hodge module with associated $\mathcal{D}_{\mathbb{G}_m}$-module $\mathcal{H}(\alpha_1,\dots,\alpha_n;\beta)$. In the sequel, we are going to compute the irregular Hodge filtration of such modules of type $(n,1)$.

Let us recall the conventions and notations used in [Castaño Domínguez and Sevenheck 2019, §4] (compare [Sabbah 2018, Notation 2.1]). We will deal with the classical hypergeometric $\mathcal{D}$-module $\mathcal{H}=\mathcal{H}(\alpha_i;\beta)$, where the α_i and β are $n+1$ real numbers belonging to the interval $[0,1)$. We will denote by $\widehat{\mathcal{H}}$ both its associated algebraic, integrable twistor $\mathcal{D}$-module on $\mathbb{G}_m$ and its underlying $\mathcal{R}^{\text{int}}_{\mathbb{A}^1_z\times\mathbb{G}_m}$-module (as in the statement of Theorem 5.7). From now on, we will write $\mathcal{X}$, ${}^\theta\mathcal{X}$ and ${}^\tau\mathcal{X}$ meaning the products $\mathbb{A}^1_z \times \mathbb{G}_{m,t}$, $\mathcal{X} \times \mathbb{G}_{m,\theta}$, and $\mathcal{X} \times \mathbb{A}^1_\tau$, respectively, where $\theta = 1/\tau$. Finally, we will write ${}^\tau\mathcal{X}_0 = \mathcal{X} \times \{\tau = 0\} \subset {}^\tau\mathcal{X}$.

Theorem 5.9. *Let real numbers $\alpha_1, \dots, \alpha_n, \beta \in [0, 1)$ be given. Suppose that $\alpha_1 \leq \cdots \leq \alpha_n$ and that moreover $\alpha_i - \beta \notin \mathbb{Z}$ for all $i = 1, \dots, n$. For each $k = 1, \dots, n$, set $\rho(k) = -(n-1)\alpha_k + k$. Then the jumping numbers of the irregular Hodge filtration of $\mathcal{H} = \mathcal{H}(\alpha_i; \beta)$ are, up to an overall real shift, the numbers $\rho(k)$. The irregular Hodge numbers are the multiplicities of those jumping numbers, or equivalently, the nonzero values of $|\rho^{-1}(x)|$, for x real.*

Moreover, for $r = 0, \dots, n-1$, let $\nu_\alpha(r) = \lceil -\alpha + r - \varepsilon - (n-1)\alpha_{r+1} \rceil$ (recall from Theorem 5.7 that $\varepsilon = \beta - \sum_{i=1}^n \alpha_i + n$). Let us consider the operators

$$\overline{Q}_r = (-(n-1))^r \prod_{i=1}^{r} (t\partial_t - \alpha_i)$$

for $r = 0, \dots, n-2$ (where the empty product equals one) and

$$\overline{Q}_{n-1} = (-(n-1))^{n-1} \prod_{i=1}^{n-1} (t\partial_t - \alpha_i) + \frac{(-(n-1))^{n-1} t(\beta - \alpha_1)}{1 + \alpha_1 - \alpha_n} \overline{Q}_0.$$

Then, the irregular Hodge filtration $F^{\mathrm{irr}}_\bullet \mathcal{H}$ is given by

$$F^{\mathrm{irr}}_{\alpha+j}\mathcal{H} = \bigoplus_{k : j \geq \nu_\alpha(k)} \mathcal{O}_X \overline{Q}_k.$$

Remark 5.10. In general, the procedure given below can be of use to find an explicit expression for the irregular Hodge filtration, not only the numbers, of any hypergeometric of type (n, m), provided both assumptions from Theorem 5.7 are fulfilled. However, the calculations become soon too cumbersome to be included here.

Proof. We will mimic the arguments of [Castaño Domínguez and Sevenheck 2019, §4], providing almost no proof of the claims which are similar to some therein.

We must first consider the rescaling of $\widehat{\mathcal{H}}$: this is the inverse image ${}^\theta\widehat{\mathcal{H}} := \mu^*\mathcal{H}$ (as $\mathcal{O}_{{}^\theta\mathcal{X}}$-module), endowed with a natural action of $\mathcal{R}^{\mathrm{int}}_{{}^\theta\mathcal{X}}$ as depicted in [Sabbah 2018, (2.4)] (note that $\theta = \tau^{-1}$), where μ is the morphism given in [ibid., Notation 2.1] by

$$\mu : {}^\theta\mathcal{X} \to \mathcal{X}$$
$$(z, t, \theta) \mapsto (z\theta, t).$$

In this sense, we can apply the same argument of [Castaño Domínguez and Sevenheck 2019, Proposition 4.1] to get that the $\mathcal{R}^{\mathrm{int}}_{{}^\theta\mathcal{X}}$-module ${}^\theta\widehat{\mathcal{H}}$ associated with $\widehat{\mathcal{H}}$ can be presented as $\mathcal{R}^{\mathrm{int}}_{{}^\theta\mathcal{X}}/(P, {}^\theta R, {}^\theta H)$, where $P = z^2\partial_z + (n-m)tz\partial_t + \varepsilon z$ as in Theorem 5.7, ${}^\theta R = z^2\partial_z - z\theta\partial_\theta$ and

$${}^\theta H = \prod_{i=1}^{n} z\theta(t\partial_t - \alpha_i) - tz\theta(t\partial_t - \beta).$$

Now we have to invert θ to obtain an $\mathcal{R}^{\text{int}}_{{}^\tau\mathcal{X}}(*{}^\tau\mathcal{X}_0)$-module ${}^\tau\widehat{\mathcal{H}}$, to work in the setting given by [Sabbah 2018, §2.3]. In this sense, we will denote by ${}^\tau\widehat{\mathcal{H}}$ the $\mathcal{R}^{\text{int}}_{{}^\tau\mathcal{X}}(*{}^\tau\mathcal{X}_0)$-module $(\mathrm{id}_\mathcal{X}\times(j\circ\mathrm{inv}))_*{}^\theta\widehat{\mathcal{H}}$, where $\mathrm{inv}:\mathbb{G}_{m,\theta}\to\mathbb{G}_{m,\tau}$ is the inversion operator $\theta\mapsto\tau^{-1}$ and $j:\mathbb{G}_{m,\tau}\hookrightarrow\mathbb{A}^1_\tau$ is the canonical inclusion. Then it is easy to see that ${}^\tau\widehat{\mathcal{H}}=\mathcal{R}^{\text{int}}_{{}^\tau\mathcal{X}}(*{}^\tau\mathcal{X}_0)/(P,{}^\tau R,{}^\tau H)$, with P as always, ${}^\tau R=z^2\partial_z+z\tau\partial_\tau$ and

$$
{}^\tau H=\prod_{i=1}^{n}\frac{z}{\tau}(t\partial_t-\alpha_i)-t\frac{z}{\tau}(t\partial_t-\beta).
$$

The next step is forming the basis of ${}^\tau\widehat{\mathcal{H}}$ as a $\mathcal{O}_{{}^\tau\mathcal{X}}(*{}^\tau\mathcal{X}_0)$-module. Let it be given by

$$
Q_k=(-(n-1))^k\prod_{i=1}^{k}\frac{z}{\tau}(t\partial_t-\alpha_i)
$$

for $k=0,\dots,n-2$ and

$$
Q_{n-1}=(-(n-1))^{n-1}\prod_{i=1}^{n-1}\frac{z}{\tau}(t\partial_t-\alpha_i)+\frac{(-(n-1))^{n-1}t(\beta-\alpha_1)}{1+\alpha_1-\alpha_n}Q_0.
$$

It is indeed a basis: we can use the expressions of ${}^\tau R$ and P to replace the classes of $z\tau\partial_\tau$ and $z^2\partial_z$, respectively, in terms of $zt\partial_t$. Now ${}^\tau\widehat{\mathcal{H}}$ is generated as a $\mathcal{O}_{{}^\tau\mathcal{X}}(*{}^\tau\mathcal{X}_0)$-module by the powers of $zt\partial_t$, and we can get rid of those of exponent greater than $n-1$ using ${}^\tau H$. The remaining n powers can be expressed as a linear combination of the Q_i, forming a triangular matrix (almost diagonal in fact), so the latter conform a basis as well.

One could wonder about the odd expression of the Q_i. In the case with no betas of [Castaño Domínguez and Sevenheck 2019], the basis considered there was formed just by the successive products $\prod_{i=1}^k\frac{z}{\tau}(t\partial_t-\alpha_i)$, up to some constant. In this case, such a basis does not provide a connection matrix solving the Birkhoff problem with a diagonal matrix as a coefficient of the pole at infinity in z, which would give us a way to read the spectrum from that matrix (see [de Gregorio et al. 2009, Proposition 4.8]). As a consequence, we have to adapt such initial basis, and that is how we get the Q_i. Let us write the connection matrix explicitly.

Let $c=(\beta-\alpha_1)/(1+\alpha_1+\alpha_n)$, in such a way that

$$
Q_{n-1}=(-(n-1))^{n-1}\prod_{i=1}^{n-1}\frac{z}{\tau}(t\partial_t-\alpha_i)+(-(n-1))^{n-1}ctQ_0.
$$

A similar (but longer) calculation to the proof of [Castaño Domínguez and Sevenheck 2019, Lemma 4.3] shows that the integrable connection arising from the $\mathcal{R}^{\text{int}}_{{}^\tau\mathcal{X}}(*{}^\tau X_0)$-module structure associated with ${}^\tau\widehat{\mathcal{H}}$ has the following matrix form:

$$
\nabla\underline{Q}=\underline{Q}\bigg((\tau A_0+zA_\infty)\frac{dz}{z^2}+(-\tau A_0+zA'_\infty)\frac{dt}{(n-1)zt}-(\tau A_0+zA_\infty)\frac{d\tau}{z\tau}\bigg).
$$

There, if $n>2$, A_0, A'_∞ and A_∞ are the matrices

$$
\begin{aligned}
A_0 &= \begin{pmatrix} 0 & \cdots & -(-(n-1))^{n-1}ct & 0 \\ 1 & \ddots & & (-(n-1))^{n-1}(c+1)t \\ & \ddots & 0 & \vdots \\ & & 1 & 0 \end{pmatrix}, \\
A'_\infty &= \operatorname{diag}((n-1)\alpha_1, \dots, (n-1)\alpha_n), \text{ and} \\
A_\infty &= \operatorname{diag}(0, 1, \dots, n-1) - \varepsilon I_n - A'_\infty.
\end{aligned}
\tag{5}
$$

If $n=2$, we have

$$
A_0 = \begin{pmatrix} ct & c(c+1)t^2 \\ 1 & (c+1)t \end{pmatrix}, \quad A'_\infty = \begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix} \quad \text{and} \quad A_\infty = \operatorname{diag}(0,1) - \varepsilon I_2 - A'_\infty. \tag{6}
$$

Finally, the irregular Hodge filtration is obtained from a suitable V-filtration along the divisor $\tau=0$ defined on ${}^\tau\widehat{\mathcal{H}}$, which is called ${}^\tau V$-filtration (the new symbol ${}^\tau V$ is to make clear the variety over which we are working; note the same convention from Remarks 2.20 on in [Sabbah 2018]). We are actually defining a filtration on ${}^\tau\widehat{\mathcal{H}}$, and then prove that it equals the ${}^\tau V$-filtration, following [Mochizuki 2015a, §2.1.2].

Let us consider

$$
\begin{aligned}
{}^\tau U_\alpha {}^\tau\widehat{\mathcal{H}} &:= \left\{ \sum_{k=0}^{n-1} f_k \tau^{\nu_k} Q_k : f_k \in \mathcal{O}_{{}^\tau\mathcal{X}}, \max(k-(n-1)\alpha_{k+1} - \varepsilon - \nu_k) \le \alpha \right\}, \\
{}^\tau U_{<\alpha} {}^\tau\widehat{\mathcal{H}} &:= \left\{ \sum_{k=0}^{n-1} f_k \tau^{\nu_k} Q_k : f_k \in \mathcal{O}_{{}^\tau\mathcal{X}}, \max(k-(n-1)\alpha_{k+1} - \varepsilon - \nu_k) < \alpha \right\},
\end{aligned}
\tag{7}
$$

for any $\alpha \in \mathbb{R}$.

The ${}^\tau U_\alpha {}^\tau\widehat{\mathcal{H}}$ form an increasing filtration, indexed by the real numbers but with a discrete set of jumping numbers, such that $\tau\, {}^\tau U_\alpha {}^\tau\widehat{\mathcal{H}} = {}^\tau U_{\alpha-1} {}^\tau\widehat{\mathcal{H}}$ for any α (those are conditions i and ii' in [Mochizuki 2015a, §2.1.2]). As usual, the graded piece associated with α is $\operatorname{Gr}_\alpha^{{}^\tau U}\, {}^\tau\widehat{\mathcal{H}} = {}^\tau U_\alpha {}^\tau\widehat{\mathcal{H}} / {}^\tau U_{<\alpha} {}^\tau\widehat{\mathcal{H}}$.

In (7), all the exponents ν_k of the powers of τ accompanying the $f_k Q_k$ satisfy that $\nu_k \ge -\alpha + k - (n-1)\alpha_{k+1} - \varepsilon$. Then we can define the steps of the filtration in the same alternative way as in [Castaño Domínguez and Sevenheck 2019, Remark 4.5] as the free $\mathcal{O}_{{}^\tau\mathcal{X}}$-modules of finite rank

$$
{}^\tau U_\alpha {}^\tau\widehat{\mathcal{H}} = \bigoplus_{k=0}^{n-1} \mathcal{O}_{{}^\tau\mathcal{X}} \cdot \tau^{\nu_\alpha(k)} Q_k, \tag{8}
$$

where $\nu_\alpha(k) = \lceil -\alpha + k - \varepsilon - (n-1)\alpha_{k+1} \rceil$. With that expression, it is clear that the graded pieces $\operatorname{Gr}_\alpha^{{}^\tau U}\, {}^\tau\widehat{\mathcal{H}}$ are

$$
\operatorname{Gr}_\alpha^{{}^\tau U}\, {}^\tau\widehat{\mathcal{H}} = \bigoplus_{k=0}^{n-1} \mathcal{O}_{\mathcal{X}} \cdot \tau^{\nu_\alpha(k)} Q_k,
$$

which are strict $\mathcal{R}_{\mathcal{X}}$-modules (condition iv in [Mochizuki 2015a, §2.1.2]).

The next step in the proof is proving that ${}^{\tau}\widehat{\mathcal{H}}$ is strictly $\mathbb{R}$-specializable along ${}^{\tau}\mathcal{X}_0$ and its ${}^{\tau}V$-filtration is actually given by the ${}^{\tau}U_\alpha{}^{\tau}\widehat{\mathcal{H}}$. Although the proof is similar to that of [Castaño Domínguez and Sevenheck 2019, Proposition 4.6], we have to adapt it a bit to our case here.

After what we already showed, it remains to show conditions iii' and v of [Mochizuki 2015a, §2.1.2] and prove that the ${}^{\tau}U_\alpha{}^{\tau}\widehat{\mathcal{H}}$ are coherent $V_0\mathcal{R}_{\mathcal{X}}$-modules. Let us start by the second condition. Consider then the mappings $\mathfrak{p}$, $\mathfrak{e}$ given by

$$(\mathfrak{p},\mathfrak{e}):\mathbb{R}\times\mathbb{C}\to\mathbb{R}\times\mathbb{C}$$
$$(\beta,\omega)\mapsto(\beta+2\Re(z\bar{\omega}),\,-\beta z+\omega-\bar{\omega}z^2).$$

We must check that the operator $z\tau\partial_\tau-\mathfrak{e}(\beta,\omega)$ is nilpotent on the graded pieces $\mathrm{Gr}_\alpha^{{}^{\tau}U}\,{}^{\tau}\widehat{\mathcal{H}}$ only for a finite amount of $(\beta,\omega)\in\mathcal{K}:=\{\beta+2\Re(z_0\bar{\omega})=\alpha\}$, for any value z_0 of z. Moreover, those (β,ω) should belong in fact to $\mathbb{R}\times\{0\}$ (see [Sabbah 2018, §1.3.a]), if we want to obtain the $\mathbb{R}$-specializability.

Take then $(\beta,\omega)\in\mathcal{K}$ and $f\tau^\nu Q_k\in{}^{\tau}U_\alpha{}^{\tau}\widehat{\mathcal{H}}$, with $f\in\mathcal{O}_{{}^{\tau}\mathcal{X}}$. We must have that $k-(n-1)\alpha_{k+1}-\varepsilon-\nu\le\alpha$. Assume that $n>2$ and $k<n-2$. Thanks to the matrix form (5) we know that

$$(z\tau\partial_\tau-\mathfrak{e}(\beta,\omega))f\tau^\nu Q_k=(z\tau\partial_\tau+(\nu+(n-1)\alpha_{k+1}+\varepsilon-k+\beta)z-\omega+\bar{\omega}z^2)(f)\tau^\nu Q_k-f\tau^{\nu+1}Q_{k+1}.$$

Recall that the α_i are increasingly ordered, lying within the interval $[0,1)$. Thus $f\tau^{\nu+1}Q_{k+1}$ lives in ${}^{\tau}U_\alpha{}^{\tau}\widehat{\mathcal{H}}$, for

$$k+1-(n-1)\alpha_{k+2}-\varepsilon-\nu-1\le((k+1)-(n-1)\alpha_{k+2}-\varepsilon)-(k-(n-1)\alpha_{k+1}-\varepsilon)-1+\alpha\le\alpha.$$

Now we should look at what happens to the class of $f\tau^{\nu+1}Q_{k+1}$ in the α-graded piece of ${}^{\tau}\widehat{\mathcal{H}}$.

Note that $[f\tau^\nu Q_k]\neq 0$ if and only if $\nu+(n-1)\alpha_{k+1}+\varepsilon-k+\alpha=0$, so

$$\begin{aligned}(z\tau\partial_\tau-\mathfrak{e}(\beta,\omega))f\tau^\nu Q_k&=(z\tau\partial_\tau+(\beta-\alpha)z-\omega+\bar{\omega}z^2)(f)\tau^\nu Q_k-f\tau^{\nu+1}Q_{k+1}\\&=(z\tau\partial_\tau-2\Re(z_0\bar{\omega})z-\omega+\bar{\omega}z^2)(f)\tau^\nu Q_k-f\tau^{\nu+1}Q_{k+1}.\end{aligned}$$

Now notice that τ divides $\tau\partial_\tau(f)$, so in fact $z\tau\partial_\tau(f)\tau^\nu Q_k\in{}^{\tau}U_{\alpha-1}{}^{\tau}\widehat{\mathcal{H}}$ and then we can further reduce our expression to

$$(z\tau\partial_\tau-\mathfrak{e}(\beta,\omega))f\tau^\nu Q_k=(-\omega-2\Re(z_0\bar{\omega})z+\bar{\omega}z^2)f\tau^\nu Q_k-f\tau^{\nu+1}Q_{k+1}.$$

On the other hand, $\tau^{\nu+1}Q_{k+1}$ does not vanish either in $\mathrm{Gr}_\alpha^{{}^{\tau}U}\,{}^{\tau}\widehat{\mathcal{H}}$ if and only if $\alpha_{k+2}=\alpha_{k+1}$. Indeed, we know that $\nu+(n-1)\alpha_{k+1}+\varepsilon-k+\alpha=0$, so doing the same as before, $k+1-(n-1)\alpha_{k+2}-\varepsilon-\nu-1=\alpha+(n-1)(\alpha_{k+2}-\alpha_{k+1})$ and the claim follows. Furthermore, in order to $(z\tau\partial_\tau-\mathfrak{e}(\beta,\omega))$ to vanish, we should impose that $\omega=0$, just by looking at the coefficients of the powers of z in the expression for f.

If $k=n-2$, we obtain from (5) that

$$\begin{aligned}&(z\tau\partial_\tau-\mathfrak{e}(\beta,\omega))f\tau^\nu Q_{n-2}\\&\quad=(z\tau\partial_\tau+(\nu+(n-1)\alpha_{n-1}+\varepsilon-(n-2)+\beta)z-\omega+\bar{\omega}z^2)(f)\tau^\nu Q_{n-2}\\&\qquad\qquad-f\tau^{\nu+1}Q_{n-1}+f\tau^{\nu+1}(-(n-1))^{n-1}ctQ_0.\end{aligned}$$

Since $-(n-1)\alpha_1-\varepsilon-\nu-1\leq-(n-1)(\alpha_1-\alpha_{n-1}+1)+\alpha<\alpha$ because $\alpha_{n-1}<\alpha_1+1$, the last summand above belongs to ${}^\tau U_{<\alpha}{}^\tau\widehat{\mathcal{H}}$, and then the argument can follow as with $k<n-2$.

Now if $k=n-1$, then everything would be the same again as before except we get the additional summand $-f\tau^{\nu+1}Q_{k+1}$, which becomes $-f\tau^{\nu+1}(-(n-1))^{n-1}(c+1)tQ_1$, whose class vanishes in the graded piece under consideration, too. Indeed,

$$1-(n-1)\alpha_2-\varepsilon-\nu-1\leq-(n-1)(\alpha_2-\alpha_n+1)+\alpha<\alpha,$$

for $\alpha_n<\alpha_2+1$.

In conclusion, $(z\tau\partial_\tau-\mathfrak{e}(\beta,\omega))^l f\tau^\nu Q_k$ can only vanish in $\mathrm{Gr}_\alpha^{{}^\tau U}\,{}^\tau\widehat{\mathcal{H}}$ if $\alpha=\beta$ (and then $\omega=0$), and does not do so until we get to an index $k+l$ such that α_{k+l} is strictly bigger than α_k. Since there is a finite set of indexes, $(z\tau\partial_\tau-\mathfrak{e}(\beta,\omega))$ is nilpotent, of nilpotency index n at most.

When $n=2$, we notice from (6) that we have two possibilities. If $k=0$, everything is the same as with $k=n-2$ for $n>2$, and if $k=1$,

$$\begin{aligned}&(z\tau\partial_\tau-\mathfrak{e}(\beta,\omega))f\tau^\nu Q_1\\&\quad=(z\tau\partial_\tau+(\nu+\alpha_2+\varepsilon-1+\beta)z-\omega+\bar{\omega}z^2)(f)\tau^\nu Q_1+f\tau^{\nu+1}(c+1)tQ_1+f\tau^{\nu+1}c(c+1)t^2Q_0.\end{aligned}$$

Here the argument runs similarly as in the general case.

Condition iii' can be rephrased as $z\tau\partial_\tau{}^\tau U_\alpha{}^\tau\widehat{\mathcal{H}}\subseteq{}^\tau U_\alpha{}^\tau\widehat{\mathcal{H}}$, using that ${}^\tau U_\alpha{}^\tau\widehat{\mathcal{H}}=\tau\,{}^\tau U_{\alpha+1}{}^\tau\widehat{\mathcal{H}}$, and that follows essentially from the same argument used to prove condition v above. Last, since $V_0\mathcal{R}_{\mathcal{X}}=\mathcal{O}_{{}^\tau\mathcal{X}}\langle z\partial_t,z\tau\partial_\tau\rangle$, it is clear from the computations above and the alternative expression (8) for the filtration steps that they are cyclic $V_0\mathcal{R}_{\mathcal{X}}$-modules, and then coherent. Summing up and noting that all the calculations performed were in fact independent of z_0, ${}^\tau\widehat{\mathcal{H}}$ is strictly $\mathbb{R}$-specializable along ${}^\tau\mathcal{X}_0$ and the ${}^\tau U_\bullet{}^\tau\widehat{\mathcal{H}}$ form its ${}^\tau V$-filtration.

We can finally show the expression for the irregular Hodge filtration and then the irregular Hodge numbers like in [Castaño Domínguez and Sevenheck 2019, Theorem 4.7]. Since we know that $\widehat{\mathcal{H}}$ underlies an object in IrrMHM($\mathbb{G}_{m,t}$) by Theorem 5.7, we deduce by [Sabbah 2018, Definition 2.52] that $\widehat{\mathcal{H}}$ is well-rescalable (see [ibid., Definition 2.19]) and so we can apply [ibid., Definition 2.22]. After formula (8), we clearly have

$$i^*_{\tau=z}{}^\tau V_\alpha{}^\tau\widehat{\mathcal{H}}={}^\tau V_\alpha{}^\tau\widehat{\mathcal{H}}/(\tau-z){}^\tau V_\alpha{}^\tau\widehat{\mathcal{H}}=\bigoplus_k\mathcal{O}_{\mathcal{X}}z^{\nu_\alpha(k)}\overline{Q}_k,$$

which is free z-graded of finite rank. Denote by π the projection $\mathcal{X}\to\mathbb{G}_{m,t}$. Then, the z-adic filtration on $\pi^*\mathcal{H}[z^{-1}]$ induces a filtration on $i^*_{\tau=z}{}^\tau V_\alpha{}^\tau\widehat{\mathcal{H}}$, given by

$$F_r i^*_{\tau=z}{}^\tau V_\alpha{}^\tau\widehat{\mathcal{H}}:=\bigoplus_{s\leq r}\Bigg(\bigoplus_{k:\nu_\alpha(k)\leq s}\mathcal{O}_{\mathbb{G}_{m,t}}\overline{Q}_k\Bigg)z^s.$$

Then, $\mathrm{Gr}^F(i^*_{\tau=z}{}^\tau U_\alpha{}^\tau\widehat{\mathcal{H}})$ is the Rees module associated to a new good filtration $F^{\mathrm{irr}}_{\alpha+\bullet}\mathcal{H}$ on $\mathcal{H}$, which is the irregular Hodge filtration. More concretely, $F^{\mathrm{irr}}_\bullet\mathcal{H}$ is given by

$$F^{\mathrm{irr}}_{\alpha+j}\mathcal{H}=\bigoplus_{k:\nu_\alpha(k)\leq j}\mathcal{O}_{\mathbb{G}_{m,t}}\overline{Q}_k.$$

Therefore, its jumping numbers are $-\varepsilon + i - 1 - (n-1)\alpha_i$ for $i = 1, \dots, n$. Since the irregular Hodge filtration is defined up to an overall real shift, we can normalize the jumping numbers to $i - (n-1)\alpha_i$ and the irregular Hodge numbers will be their multiplicities. □

Acknowledgements

We would like to thank Takuro Mochizuki for communicating the proof of Lemma 2.3 to us. We would further like to thank Takuro Mochizuki and Claude Sabbah for their interest in our work and for many stimulating discussions. We are grateful to the participants of the workshop *Mixed Twistor D-modules* in Heidelberg in 2017 for their time and effort. We also thank the Max Planck Institute for Mathematics in the Sciences, where some part of the work presented here has been carried out.

References

[Castaño Domínguez and Sevenheck 2019] A. Castaño Domínguez and C. Sevenheck, "Irregular Hodge filtration of some confluent hypergeometric systems", *J. Inst. Math. Jussieu* (online publication May 2019).

[Esnault et al. 2017] H. Esnault, C. Sabbah, and J.-D. Yu, "E_1-degeneration of the irregular Hodge filtration", *J. Reine Angew. Math.* **729** (2017), 171–227. MR Zbl

[Fedorov 2018] R. Fedorov, "Variations of Hodge structures for hypergeometric differential operators and parabolic Higgs bundles", *Int. Math. Res. Not.* **2018**:18 (2018), 5583–5608. MR Zbl

[Gelfand et al. 1987] I. M. Gelfand, M. I. Graev, and A. V. Zelevinski, "Holonomic systems of equations and series of hypergeometric type", *Dokl. Akad. Nauk SSSR* **295**:1 (1987), 14–19. In Russian; translated in *Soviet Math. Dokl.* **36**:1 (1988), 5–10. MR Zbl

[Gelfand et al. 1989] I. M. Gelfand, A. V. Zelevinski, and M. M. Kapranov, "Гипергеометрические функции и торические многообразия" "(Hypergeometric functions and toric varieties)", *Funktsional. Anal. i Prilozhen.* **23**:2 (1989), 12–26. Translated as "Hypergeometric functions and toral manifolds", *Funct. Anal. Appl.* **23**:2 (1989), 94–106. MR Zbl

[de Gregorio et al. 2009] I. de Gregorio, D. Mond, and C. Sevenheck, "Linear free divisors and Frobenius manifolds", *Compos. Math.* **145**:5 (2009), 1305–1350. MR Zbl

[Katz 1990] N. M. Katz, *Exponential sums and differential equations*, Annals of Mathematics Studies **124**, Princeton University Press, 1990. MR Zbl

[Katzarkov et al. 2017] L. Katzarkov, M. Kontsevich, and T. Pantev, "Bogomolov–Tian–Todorov theorems for Landau–Ginzburg models", *J. Differential Geom.* **105**:1 (2017), 55–117. MR Zbl

[Matusevich et al. 2005] L. F. Matusevich, E. Miller, and U. Walther, "Homological methods for hypergeometric families", *J. Amer. Math. Soc.* **18**:4 (2005), 919–941. MR Zbl

[Mochizuki 2011] T. Mochizuki, *Wild harmonic bundles and wild pure twistor D-modules*, Astérisque **340**, Société Mathématique de France, Paris, 2011. MR Zbl

[Mochizuki 2015a] T. Mochizuki, *Mixed twistor $\mathcal{D}$-modules*, Lecture Notes in Mathematics **2125**, Springer, 2015. MR Zbl

[Mochizuki 2015b] T. Mochizuki, "Twistor property of GKZ-hypergeometric systems", preprint, 2015. arXiv

[Reichelt 2014] T. Reichelt, "Laurent polynomials, GKZ-hypergeometric systems and mixed Hodge modules", *Compos. Math.* **150**:6 (2014), 911–941. MR Zbl

[Reichelt and Sevenheck 2015] T. Reichelt and C. Sevenheck, "Hypergeometric Hodge modules", preprint, 2015. To appear in *Algebraic Geometry*. arXiv

[Sabbah 2018] C. Sabbah, *Irregular Hodge theory*, Mém. Soc. Math. Fr. (N.S.) **156**, Société Mathématique de France, Paris, 2018. MR Zbl

[Sabbah and Schnell 2018] C. Sabbah and C. Schnell, "The MHM project", electronic reference, 2018, Available at http://www.cmls.polytechnique.fr/perso/sabbah/MHMProject/mhm.html.

[Sabbah and Yu 2015] C. Sabbah and J.-D. Yu, "On the irregular Hodge filtration of exponentially twisted mixed Hodge modules", *Forum Math. Sigma* **3** (2015), art. id. e9, 71 pp. MR Zbl

[Schulze and Walther 2009] M. Schulze and U. Walther, "Hypergeometric D-modules and twisted Gauß-Manin systems", *J. Algebra* **322**:9 (2009), 3392–3409. MR Zbl

[Serre 1955–1956] J.-P. Serre, "Géométrie algébrique et géométrie analytique", *Ann. Inst. Fourier, Grenoble* **6** (1955–1956), 1–42. MR Zbl

[Yu 2014] J.-D. Yu, "Irregular Hodge filtration on twisted de Rham cohomology", *Manuscripta Math.* **144**:1-2 (2014), 99–133. MR Zbl

Communicated by Hélène Esnault
Received 2018-07-17 Revised 2019-01-28 Accepted 2019-03-08

alberto.castano@usc.es *Instituto de Matemáticas, Universidade de Santiago de Compostela, Spain*

Fakultät für Mathematik, Technische Universität Chemnitz, Chemnitz, Germany

treichelt@mathi.uni-heidelberg.de *Mathematisches Institut, Universität Heidelberg, Heidelberg, Germany*

christian.sevenheck@mathematik.tu-chemnitz.de *Fakultät für Mathematik, Technische Universität Chemnitz, Chemnitz, Germany*

ALGEBRA AND NUMBER THEORY 13:6 (2019)
dx.doi.org/10.2140/ant.2019.13.1443

Ulrich bundles on K3 surfaces

Daniele Faenzi

We show that any polarized K3 surface supports special Ulrich bundles of rank 2.

Given an n-dimensional closed subvariety $X \subset \mathbb{P}^N$, a coherent sheaf $\mathcal{F}$ on X is Ulrich if $\mathrm{H}^*(\mathcal{F}(-t)) = 0$ for $1 \le t \le n$. We refer to [Coskun 2017; Beauville 2018] for an introduction. We mention that Ulrich sheaves are related to Chow forms (this was their main motivation for the study in [Eisenbud et al. 2003]), to determinantal representations and generalized Clifford algebras, to Boij–Söderberg theory [Schreyer and Eisenbud 2010], to the minimal resolution conjecture, and to the representation type of varieties [Faenzi and Pons-Llopis 2015].

Conjecturally, Ulrich sheaves exist for any X, see [Eisenbud et al. 2003]. They are known to exist for several classes of varieties e.g., complete intersections, curves, Veronese, Segre, Grassmann varieties. Low-rank Ulrich bundles on surfaces have been studied intensively, and Ulrich bundles of rank 2 (or sometimes 1) are known in many cases. We refer to [Casnati 2017; Beauville 2018] for a survey and further references. Let us only review some of the cases that are most relevant for us, namely among surfaces with trivial canonical bundle.

In [Beauville 2016], Ulrich bundles of rank 2 are proved to exist on abelian surfaces. In [Aprodu et al. 2017], it is proved that K3 surfaces support Ulrich bundles of rank 2, provided that some Noether–Lefschetz open condition is satisfied. The case of quartic surfaces was previously analyzed in detail in [Coskun et al. 2012]. The main techniques used so far are the Serre construction starting from points on X and Lazarsfeld–Mukai bundles.

In this note, we show that any K3 surface supports an Ulrich bundle $\mathcal{E}$ of rank 2 with $c_1(\mathcal{E}) = 3H$, for any polarization H. So these bundles are *special* [Eisenbud et al. 2003]. We allow singular surfaces with trivial canonical bundle. The main tool is an enhancement of Serre's construction based on unobstructedness of simple sheaves on a $K3$ surface.

Let us state the result more precisely. We work over an algebraically closed field $\boldsymbol{k}$. Let X be an integral (i.e., reduced and irreducible) projective surface with $\omega_X \simeq \mathcal{O}_X$ and $\mathrm{H}^1(\mathcal{O}_X) = 0$. We denote by X_{sm} the smooth locus of X.

Fix a very ample divisor H on X. Under the closed embedding given by the complete linear series $|\mathcal{O}_X(H)|$ we may view X as a subvariety of some projective space $\mathbb{P}^g$. A hyperplane section C of X

Author partially supported by ISITE-BFC project (contract ANR-lS-IDEX-OOOB).
MSC2010: primary 14F05; secondary 13C14, 14J60.
Keywords: ACM vector sheaves and bundles, Ulrich sheaves, K3 surfaces.

is a projective Gorenstein curve of arithmetic genus g with $\omega_C \simeq \mathcal{O}_C(H)$, where H also denotes the restriction of H to C. We may choose C to be integral too.

A locally Cohen–Macaulay sheaf $\mathcal{E}$ on X is arithmetically Cohen–Macaulay (ACM) if $\mathrm{H}^1(\mathcal{E}(tH)) = 0$ for all $t \in \mathbb{Z}$. A special class of ACM sheaves are Ulrich sheaves, which are characterized by the property $\mathrm{H}^*(\mathcal{E}(-tH)) = 0$ for $t = 1, 2$. Of course all these notions depend on the polarization H. We call simple a sheaf whose only endomorphisms are homotheties.

Theorem 1. *Let X and H be as above. Then there exists a simple Ulrich vector bundle of rank 2 on X whose determinant is $\mathcal{O}_X(3H)$.*

The strategy to prove the theorem is the following. First we build an ACM vector bundle $\mathcal{E}$ of rank 2 by Serre's construction applied to a projective coordinate system in X. Then we perform an elementary modification of $\mathcal{E}$ along a single generic point $p \in X$, producing a simple nonreflexive sheaf having the Chern character of an Ulrich bundle. Finally we flatly deform such sheaf and check that generically this yields the desired Ulrich bundle.

Prior to all this, we start by observing that the trivial bundle is a (trivial) example of ACM line bundle. Indeed, using that $\mathrm{H}^1(\mathcal{O}_X) = 0$ and that C is connected, one checks that $\mathrm{H}^1(\mathcal{O}_X(-H)) = 0$. In turn, this easily implies $\mathrm{H}^1(\mathcal{O}_X(-tH)) = 0$ for all $t \geq 2$. Also, Serre duality and triviality of ω_X give $\mathrm{H}^1(\mathcal{O}_X(tH)) = 0$ for all $t \geq 0$. This way, we see that $\mathcal{O}_X$ is an ACM line bundle on X. Combining this with Max Noether's theorem on the generation of the canonical ring of curves (see [Rosenlicht 1952] for a version for Gorenstein curves) one obtains, working as in [Saint-Donat 1974, Theorem 6.1], that $X \subset \mathbb{P}^g$ is an ACM surface of degree $2g - 2$.

However this line bundle is never Ulrich, nor is any line bundle of the form $\mathcal{O}_X(dH)$. So generically (for instance when X has Picard number 1) the surface X will not support Ulrich line bundles. We thus move to rank two and start by constructing a simple ACM bundle.

Lemma 2. *Let $Z \subset X_{\mathrm{sm}}$ be a set of $g+2$ points in general linear position. Then there is a unique coherent sheaf $\mathcal{E}$ of rank 2 fitting into a nonsplitting exact sequence:*

$$0 \to \mathcal{O}_X \to \mathcal{E} \to \mathcal{I}_Z(H) \to 0. \tag{1}$$

The sheaf $\mathcal{E}$ is locally free, simple and ACM. It satisfies

$$\mathcal{E} \simeq \mathcal{E}^*(H), \quad \mathrm{h}^0(\mathcal{E}) = 1, \quad \mathrm{h}^1(\mathcal{E}) = \mathrm{h}^2(\mathcal{E}) = 0, \quad \mathrm{ext}^1_X(\mathcal{E}, \mathcal{E}) = 2g + 4.$$

Proof. Taking cohomology of the exact sequence

$$0 \to \mathcal{I}_Z(H) \to \mathcal{O}_X(H) \to \mathcal{O}_Z \to 0, \tag{2}$$

and using the fact that Z is in general linear position and hence contained in no hyperplane, we get $\mathrm{H}^0(\mathcal{I}_Z(H)) = 0$ and $\mathrm{h}^1(\mathcal{I}_Z(H)) = 1$.

By Serre duality we get $\mathrm{ext}^1_X(\mathcal{I}_Z(H), \mathcal{O}_X) = \mathrm{h}^1(\mathcal{I}_Z(H)) = 1$ so, up to proportionality, there is a unique nonsplitting extension of the desired form. Correspondingly, there exists a unique coherent sheaf $\mathcal{E}$ of rank two fitting into a nonsplitting exact sequence of the form (1). The sheaf $\mathcal{E}$ we obtain this way satisfies $h^0(\mathcal{E}) = 1$ and $H^1(\mathcal{E}) \simeq \mathrm{Ext}^1_X(\mathcal{E}, \mathcal{O}_X)^* = 0$ because applying $\mathrm{Hom}_X(-, \mathcal{O}_X)$ to (1) we obtain a nonzero map (and thus an isomorphism) $H^0(\mathcal{O}_X) \to \mathrm{Ext}^1_X(\mathcal{I}_Z(H), \mathcal{O}_X)$.

This map is the dual of the homomorphism $H^1(\mathcal{I}_Z(H)) \to H^2(\mathcal{O}_X)$ obtained by taking global sections in (1). So $H^1(\mathcal{E}) = H^2(\mathcal{E}) = 0$.

If X is smooth we deduce that $\mathcal{E}$ is locally free from the Cayley–Bacharach property, see for instance [Huybrechts and Lehn 1997, Theorem 5.1.1]. Indeed, since Z is in general linear position (i.e., Z is a projective frame in $\mathbb{P}^g$), no hyperplane passes through any subset of $g+1$ points of Z. Anyway the statement follows in general by a minor modification of the argument appearing in [Faenzi and Pons-Llopis 2015, Lemma 7.2]. Indeed by the local-to-global spectral sequence, using $H^1(\mathcal{O}_X(-H)) = 0$ and $\mathcal{H}om_X(\mathcal{I}_Z(H), \mathcal{O}_X) \simeq \mathcal{O}_X(-H)$ we get the following exact sequence:

$$0 \to \mathrm{Ext}^1_X(\mathcal{I}_Z(H), \mathcal{O}_X) \to H^0(\mathcal{E}xt^1_X(\mathcal{I}_Z(H), \mathcal{O}_X)) \to H^2(X, \mathcal{O}_X(-H)) \to 0.$$

In turn, using $\mathcal{E}xt^1_X(\mathcal{I}_Z(H), \mathcal{O}_X) \simeq \omega_Z \simeq \mathcal{O}_Z$ and $H^2(X, \mathcal{O}_X(-H)) \simeq H^0(X, \mathcal{O}_X(H))^*$, if we choose Z to be a projective coordinate system of $\mathbb{P}^g$, we rewrite this exact sequence as

$$0 \to \mathrm{Ext}^1_X(\mathcal{I}_Z(H), \mathcal{O}_X) \to H^0(\mathcal{O}_Z) \xrightarrow{M} H^0(X, \mathcal{O}_X(H))^* \to 0,$$

where

$$M = \begin{pmatrix} 1 & \cdots & 0 & 1 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 1 \end{pmatrix}.$$

So $\mathrm{Ext}^1_X(\mathcal{I}_Z(H), \mathcal{O}_X)$ is generated by the vector $(1, \dots, 1, -1)^t$ and since this vector corresponds to an extension in $\mathcal{E}xt^1_X(\mathcal{I}_Z(H), \mathcal{O}_X)$ which is nonzero at any point of Z we have that the sequence defining $\mathcal{E}$ is locally nonsplit around each point of Z, which in turn implies that $\mathcal{E}$ is locally free at each such point (and hence everywhere). From $c_1(\mathcal{E}) = H$, since $\mathcal{E}$ is locally free of rank 2, we get a canonical isomorphism $\mathcal{E} \simeq \mathcal{E}^*(H)$.

Let us prove that $\mathcal{E}$ is ACM. We already have $h^1(\mathcal{E}) = 0$ and thus by Serre duality $h^1(\mathcal{E}(-H)) = h^1(\mathcal{E}^*(H)) = h^1(\mathcal{E}) = 0$. Also $h^0(\mathcal{E}(-H)) = 0$ and $h^2(\mathcal{E}(-H)) = 1$. Note that, choosing an integral hyperplane section curve C that avoids Z, (1) becomes:

$$0 \to \mathcal{O}_C \to \mathcal{E}|_C \to \mathcal{O}_C(H) \to 0.$$

From $H^k(\mathcal{E}(-H)) = 0$ for $k = 0, 1$ we deduce $h^0(\mathcal{E}|_C) = 1$ so the previous exact sequence does not split. Then $h^0(\mathcal{E}|_C(-H)) = 0$. This easily implies $H^1(\mathcal{E}(-2H)) = 0$ and actually $H^1(\mathcal{E}(-tH)) = 0$ for all $t \geq 2$. Serre duality now gives $H^1(\mathcal{E}(tH)) = 0$ for all $t \geq 1$. In other words $\mathcal{E}$ is ACM.

It remains to check that $\mathcal{E}$ is simple. Applying $\mathrm{Hom}_X(\mathcal{E}, -)$ to the exact sequence (2) we get that the nonzero space $\mathrm{Hom}_X(\mathcal{E}, \mathcal{I}_Z(H))$ is contained in $\mathrm{Hom}_X(\mathcal{E}, \mathcal{O}_X(H)) \simeq H^0(\mathcal{E}) \simeq \boldsymbol{k}$, so $\hom_X(\mathcal{E}, \mathcal{I}_Z(H)) = 1$. As $\mathcal{H}om_X(\mathcal{E}, \mathcal{O}_Z)$ is a skyscraper sheaf of rank 2 at Z we have $\mathrm{ext}^k_X(\mathcal{E}, \mathcal{O}_Z) = (2g+4)\delta_{0,k}$. We deduce $\mathrm{ext}^1_X(\mathcal{E}, \mathcal{I}_Z(H)) = 2g+4$ and $\mathrm{ext}^0_X(\mathcal{E}, \mathcal{I}_Z(H)) = 0$.

Therefore, applying $\mathrm{Hom}_X(\mathcal{E},-)$ to the (1), since $\mathrm{Hom}_X(\mathcal{E},\mathcal{O}_X) \simeq \mathrm{h}^2(\mathcal{E}) = 0$ we get that $\mathrm{End}_X(\mathcal{E})$ is contained in $\mathrm{Hom}_X(\mathcal{E},\mathcal{I}_Z(H))$ and is therefore 1-dimensional. This says that $\mathcal{E}$ is simple. By Serre duality $\mathrm{ext}^2_X(\mathcal{E},\mathcal{E}) = 1$. We deduce $\mathrm{ext}^1_X(\mathcal{E},\mathcal{E}) = \mathrm{ext}^1_X(\mathcal{E},\mathcal{I}_Z(H)) = 2g+4$. □

Given a reduced subscheme $Z \in \mathrm{Hilb}_{g+2}(X_{\mathrm{sm}})$ consisting of points in general linear position, there is a unique rank-2 bundle associated with Z according to the previous lemma. We denote it by $\mathcal{E}_Z$. We write $\mathcal{O}_p$ for the skyscraper sheaf of a point $p \in X$.

Lemma 3. *Assume $\eta : \mathcal{E}_Z \to \mathcal{O}_p$ is surjective. Then $\mathcal{E}^\eta = \ker(\eta)$ is a simple sheaf with*

$$c_1(\mathcal{E}^\eta) = H, \quad c_2(\mathcal{E}^\eta) = g+3, \quad \mathrm{ext}^1_X(\mathcal{E}^\eta,\mathcal{E}^\eta) = 2g+8.$$

Proof. Recall that $\mathcal{E} = \mathcal{E}_Z$ is simple and observe that this implies $\mathrm{Hom}_X(\mathcal{E},\mathcal{E}^\eta) = 0$, as the composition of any nonzero map $\mathcal{E} \to \mathcal{E}^\eta$ with $\mathcal{E}^\eta \hookrightarrow \mathcal{E}$ would provide a self-map of $\mathcal{E}$ which is not a multiple of the identity. Also, since $\mathcal{E}$ is locally free we have $\mathrm{hom}_X(\mathcal{E},\mathcal{O}_p) = 2$ and $\mathrm{Ext}^k_X(\mathcal{E},\mathcal{O}_p) = 0$ for $k > 0$. Therefore, using Lemma 2 and applying $\mathrm{Hom}_X(\mathcal{E},-)$ to the exact sequence:

$$0 \to \mathcal{E}^\eta \to \mathcal{E} \to \mathcal{O}_p \to 0. \tag{3}$$

we obtain $\mathrm{ext}^1_X(\mathcal{E},\mathcal{E}^\eta) = 2g+5$ and $\mathrm{ext}^2_X(\mathcal{E},\mathcal{E}^\eta) = 1$.

Next, Serre duality gives $\mathrm{ext}^k_X(\mathcal{O}_p,\mathcal{E}) = 2\delta_{2,k}$, while $\mathrm{ext}^k_X(\mathcal{O}_p,\mathcal{O}_p)$ is the dimension of the k-th exterior power of the normal bundle of p in X and thus takes value $\binom{2}{k}$. Therefore, applying $\mathrm{Hom}_X(\mathcal{O}_p,-)$ to (3) we find $\mathrm{ext}^1_X(\mathcal{O}_p,\mathcal{E}^\eta) = 1$ and $\mathrm{ext}^2_X(\mathcal{O}_p,\mathcal{E}^\eta) = 3$. Putting these computations together and applying

$$\mathrm{hom}_X(\mathcal{E}^\eta,\mathcal{E}^\eta) = \mathrm{ext}^2_X(\mathcal{E}^\eta,\mathcal{E}^\eta) = 1, \quad \mathrm{ext}^1_X(\mathcal{E}^\eta,\mathcal{E}^\eta) = 2g+8.$$

The computation of Chern classes is straightforward. □

Lemma 4. *Let $p \in X_{\mathrm{sm}} \setminus Z$. Then, for a generic map $\eta : \mathcal{E}_Z \to \mathcal{O}_p$, the induced map on global sections $\mathrm{H}^0(\eta) : \mathrm{H}^0(\mathcal{E}_Z) \to \mathrm{H}^0(\mathcal{O}_p)$ is an isomorphism.*

Proof. Put $\mathcal{E} = \mathcal{E}_Z$. It suffices to check that there exists η such that the induced map $\mathrm{H}^0(\eta) : \boldsymbol{k} \simeq \mathrm{H}^0(\mathcal{E}) \to \mathrm{H}^0(\mathcal{O}_p) \simeq \boldsymbol{k}$ is an isomorphism, for this is an open condition. To do it, we apply $\mathrm{Hom}_X(\mathcal{I}_Z(H),-)$ to the exact sequence:

$$0 \to \mathcal{I}_p \to \mathcal{O}_X \to \mathcal{O}_p \to 0.$$

This gives an exact sequence:

$$\mathrm{Ext}^1_X(\mathcal{I}_Z(H),\mathcal{I}_p) \to \mathrm{Ext}^1_X(\mathcal{I}_Z(H),\mathcal{O}_X) \to \mathrm{Ext}^1_X(\mathcal{I}_Z(H),\mathcal{O}_p).$$

Observe that $\mathcal{H}om_X(\mathcal{I}_Z(H),\mathcal{O}_p) \simeq \mathcal{O}_p$ and $\mathcal{E}xt^1_X(\mathcal{I}_Z(H),\mathcal{O}_p) = 0$ as these sheaves are computed locally on X and, since $p \cap Z = \varnothing$, we may choose an open cover of X consisting of subsets where $\mathcal{I}_Z$ is trivial or $\mathcal{O}_p$ vanishes. Then the local-to-global spectral sequence gives $\mathrm{Ext}^1_X(\mathcal{I}_Z(H),\mathcal{O}_p) = 0$ so the extension corresponding to (1) admits a lifting to $\mathcal{I}_p$. In other words, we get the commutative exact diagram:

$$
\begin{array}{ccccccccc}
 & & 0 & & 0 & & & & \\
 & & \downarrow & & \downarrow & & & & \\
0 & \longrightarrow & \mathcal{I}_p & \longrightarrow & \mathcal{E}^\eta & \longrightarrow & \mathcal{I}_Z(H) & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow & & \| & & \\
0 & \longrightarrow & \mathcal{O}_X & \longrightarrow & \mathcal{E} & \longrightarrow & \mathcal{I}_Z(H) & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow{\scriptstyle \eta} & & & & \\
 & & \mathcal{O}_p & = & \mathcal{O}_p & & & & \\
 & & \downarrow & & \downarrow & & & & \\
 & & 0 & & 0 & & & &
\end{array}
$$

where η and $\mathcal{E}^\eta$ are defined by the diagram. For this choice of η we get, by the top row of the diagram, $\mathrm{H}^0(\mathcal{E}^\eta)=0$, which implies that $\mathrm{H}^0(\eta)$ is an isomorphism. □

By the previous lemma, we may choose $\mathcal{E}_Z$ as in Lemma 2, a point $p \in X_{\mathrm{sm}} \setminus Z$, some $\eta : \mathcal{E}_Z \twoheadrightarrow \mathcal{O}_p$ and consider the sheaf $\mathcal{E}^\eta$. The goal is to deform $\mathcal{E}^\eta(H)$ to an Ulrich bundle. We use the notation $\mathcal{F}_s^*$ for $(\mathcal{F}_s)^*$ (which is a priori not the same as $(\mathcal{F}^*)_s$).

Lemma 5. *There exist a smooth connected variety S_0 of dimension $2g+8$ and a flat family of simple sheaves $\mathcal{F}$ on $X \times S_0$ such that $\mathcal{F}_s(H)$ is an Ulrich bundle for s generic in S_0 and $\mathcal{F}_{s_0} \simeq \mathcal{E}^\eta$ for some distinguished point s_0 of S_0.*

Proof. We proved in Lemma 3 that $\mathcal{E}^\eta$ is simple. Since the nonlocally free locus of $\mathcal{E}^\eta$ is disjoint from the singular locus of X, we may apply the arguments of [Mukai 1984, Theorem 0.1]. In particular [Altman and Kleiman 1980] the moduli functor of simple sheaves on X is prorepresented by a moduli space Spl_X which can be constructed in the étale topology and which is smooth of dimension $2g+8$ at $\mathcal{E}^\eta$ (this is essentially [Mukai 1984, Theorem 0.3]). Therefore there exists an open piece of Spl_X which is a quasiprojective variety S equipped with a flat family $\mathcal{F}$ of simple sheaves on X, such that the induced map $S \to \mathrm{Spl}_X$ is a local isomorphism around the point corresponding to $\mathcal{E}^\eta$. We denote this point by s_0, so that $\mathcal{F}_{s_0} \simeq \mathcal{E}^\eta$.

We may assume that S is smooth and connected of dimension $2g+8$. Since the reflexive hull $\mathcal{E}$ of $\mathcal{E}^\eta$ is locally free and satisfies the assumption of [Artamkin 1990, Corollary 1.5], we get that $\mathcal{F}_s$ is locally free for all s in an open dense subset S_1 of S.

Now observe that $\mathrm{H}^*(\mathcal{F}_{s_0})=0$ by Lemmas 2 and 4. Then, semicontinuity ensures that $\mathrm{H}^*(\mathcal{F}_s)=0$ for all s in an open dense subset S_0 of S_1. Therefore, the isomorphism $\mathcal{F}_s^* \simeq \mathcal{F}_s(-H)$ and Serre duality give $\mathrm{H}^i(\mathcal{F}_s(-H)) \simeq \mathrm{H}^{2-i}(\mathcal{F}_s^*(H))^* \simeq \mathrm{H}^{2-i}(\mathcal{F}_s)^* = 0$. This says that $\mathcal{F}_s(H)$ is a special Ulrich bundle, for all $s \in S_0$. □

For the reader's benefit we also provide a proof of Lemma 5 independent of [Artamkin 1990]. The point is to check that $\mathcal{F}_s$ is locally free for all s in an open dense subset of S. To do this, first recall again that the nonlocally free locus of $\mathcal{E}^\eta$ is disjoint from the singular locus of X, so up to shrinking S we may assume that this happens for $\mathcal{F}_s$ for all $s \in S$. Then $\mathcal{F}_s^{**}$ is locally free for $s \in S$.

Next, we may find an integer $t_0 \leq -1$ such that $\mathrm{H}^0(\mathcal{F}_s^{**}(t_0 H)) = \mathrm{H}^1(\mathcal{F}_s^{**}(t_0 H)) = 0$ for all $s \in S$. This can be done for instance using Kollar's theory of husks [2008], which gives a stratification $(S_i)_{i=1,\dots,r}$ of S such that $\mathcal{F}_s^{**}$ defines a flat family of sheaves on X parametrized by S_i. Using base change over each S_i one finds t_i satisfying the required vanishing together with $\mathrm{H}^0(\mathcal{F}_s^{**}(t_i H)|_C) = 0$, for a fixed curve $C \in |\mathcal{O}_X(H)|$. Then t_0 can be taken to be the minimum among $t_1, \dots, t_r$.

Recall that $\mathrm{H}^*(\mathcal{F}_{s_0}) = 0$ and observe that (3) gives:

$$\mathrm{h}^1(\mathcal{F}_{s_0}(tH)) = \begin{cases} 1 & \text{if } t \leq -1, \\ 0 & \text{if } t \geq 0. \end{cases}$$

By semicontinuity, we have that $\mathrm{H}^*(\mathcal{F}_s) = 0$, $\mathrm{h}^1(\mathcal{F}_s(tH)) = 0$ for all $t \geq 0$ and $\mathrm{h}^1(\mathcal{F}_s(tH)) \leq 1$ for $t \leq -1$ for all s in an open dense subset of S. We still call S this subset.

Next, for all $s \in S$ we consider the double dual sequence

$$0 \to \mathcal{F}_s \to \mathcal{F}_s^{**} \to \tau(\mathcal{F}_s) \to 0, \tag{4}$$

where the torsion sheaf $\tau(\mathcal{F}_P)$ is defined by the sequence. Put ℓ_s for the length of $\tau(\mathcal{F}_s)$.

Since $\mathrm{H}^0(\mathcal{F}_s^{**}(t_0 H)) = \mathrm{H}^1(\mathcal{F}_s^{**}(t_0 H)) = 0$, from the previous exact sequence we get $\ell_s = \mathrm{h}^0(\tau(\mathcal{F}_s)) = \mathrm{h}^1(\mathcal{F}_s(t_0 H)) \leq 1$ (we neglect to indicate the twist on zero-dimensional sheaves).

Now we have two alternatives. Namely, either for s general enough in S one has $\ell_s = 0$, i.e., $\tau(\mathcal{F}_s) = 0$; or otherwise for all $s \in S$ we get $\ell_s = 1$, i.e., $\tau(\mathcal{F}_s) \simeq \mathcal{O}_{p_s}$, for some point $p_s \in X$ with $p_{s_0} = p$.

In the first case, we have $\mathcal{F}_s \simeq \mathcal{F}_s^{**}$ and $\mathcal{F}_s$ is locally free. So we would like to rule out the second alternative. By contradiction we assume that, for all $s \in S$, we have $\tau(\mathcal{F}_s) \simeq \mathcal{O}_{p_s}$. This gives a map $\gamma : S \to X$ associating p_s to s. This time $\mathcal{F}^{**}$ is flat over S and (4) is the restriction to $X \times \{s\}$ of a sequence on $X \times S$:

$$0 \to \mathcal{F} \to \mathcal{F}^{**} \to \tau(\mathcal{F}) \to 0,$$

with $(\mathcal{F}_s)^{**} \simeq (\mathcal{F}^{**})_s$ and where $\tau(\mathcal{F})$ is a line bundle supported on the graph of γ.

Also, again the previous exact sequence together with $\mathrm{H}^*(\mathcal{F}_s) = 0$ gives $\mathrm{h}^0(\mathcal{F}_s^{**}) = 1$ so that $\mathcal{F}_s^{**}$ has a unique nonzero global section up to a scalar. This section vanishes along a subscheme $Z_s \subset X$ and, up to shrinking again S we may assume that Z_s is zero-dimensional reduced and in general linear position, because these are open conditions, so that $\mathcal{F}_s^{**} \simeq \mathcal{E}_{Z_s}$.

For each sheaf $\mathcal{F}_s^{**}$ of this family, we denote by $\eta_s : \mathcal{F}_s^{**} \twoheadrightarrow \mathcal{O}_{p_s}$ the induced surjection of $\mathcal{F}_s^{**}$ onto $\tau(\mathcal{F}_s)$. We think of η_s as an element of $\mathbb{P}(\mathrm{H}^0(\mathcal{F}_s^{**}|_{p_s})) \simeq \mathbb{P}^1$ (we adopt the convention of writing $\mathbb{P}(V)$ for the projective space of hyperplanes of a vector space V). Plainly, we have $\mathcal{F}_{s_0}^{**} \simeq \mathcal{E}^\eta$, $\tau(\mathcal{F}_{s_0}) \simeq \mathcal{O}_p$ and η_{s_0} is identified with η. Note that $\mathcal{F}_s = \ker(\eta_s)$.

We assert that the family $\mathcal{F}$ is parametrized by an open subset T of the set of triples:

$$\{(W, q, \xi) \mid W \in \mathrm{Hilb}_{g+2}(X),\ q \in X,\ \xi \in \mathbb{P}(\mathrm{H}^0(\mathcal{E}_W|_q))\}.$$

The subset T consists of triples (W, q, ξ) with $W \subset X_{\mathrm{sm}}$ reduced and in general linear position in X, $q \in X_{\mathrm{sm}} \setminus W$ and ξ is surjective. Given such a triple, we get that the sheaf $\ker(\xi)$ is simple by Lemma 3. Clearly this gives a flat deformation of $\mathcal{E}^\eta$ so, because $S \to \mathrm{Spl}_X$ is a local isomorphism at $\mathcal{E}^\eta$, there is a possibly smaller open subset T_0 such that all the resulting sheaves $\ker(\xi)$ are of the form $\mathcal{F}_s$, for some $s \in S$. By construction any sheaf $\mathcal{F}_s$ should be of this form by taking $q = p_s$, $W = Z_s$ and $\xi = \eta_s$.

But T_0 is an open dense subset of a $\mathbb{P}^1$-bundle over an open subset of $\mathrm{Hilb}_{g+2}(X) \times X$ and thus has dimension $1 + 2(g + 2) + 2 = 2g + 7$. Therefore T_0 cannot dominate S, as $\dim(S) = 2g + 8$. This says that the second alternative does not take place, so we have proved that $\mathcal{F}_s(H)$ is an Ulrich bundle for general s.

Recall the notation $\mathrm{M}_X(v)$ for the moduli space of H-semistable sheaves $\mathcal{F}$ on X whose Mukai vector $v = (v_0, v_1, v_2)$ satisfies $v_0 = \mathrm{rk}(\mathcal{F})$, $v_1 = c_1(\mathcal{F})$ and $v_2 = \chi(\mathcal{F}) - \mathrm{rk}(\mathcal{F})$. From [Qin 1993, Lemma 2.1] we obtain the following stronger version of Theorem 1.

Corollary 6. *If X is smooth, $\mathrm{M}_X(2, H, -2)$ is of dimension $2g + 8$ and a general point of it corresponds to a sheaf $\mathcal{E}$ which is stable (with respect to all polarizations) and such that $\mathcal{E}(H)$ is a special Ulrich bundle.*

Again, we also offer a proof independent of [Qin 1993; Artamkin 1990]. Consider the family of Ulrich sheaves $\mathcal{F}(H)$ with parameter space S_0 constructed in the previous lemma. Recall that, for generic $s \in S_0$, the sheaf $\mathcal{F}_s(H)$ is Ulrich, hence semistable with Ulrich sheaves as Jordan–Hölder factors [Faenzi and Pons-Llopis 2015, Lemma 7.1]. So we have to check that $\mathcal{F}_s$ is not strictly semistable. If it was, we would have an exact sequence:

$$0 \to \mathcal{L} \to \mathcal{F}_s \to \mathcal{L}^*(H) \to 0, \tag{5}$$

where $\mathcal{L}(H)$ is an Ulrich sheaf or rank 1 on X. Actually $\mathcal{L}(H)$ is an Ulrich line bundle since X is smooth. Since $\mathcal{L}$ and $\mathcal{L}^*(H)$ are rigid in view of $\mathrm{H}^1(\mathcal{O}_X) = 0$, they do not depend on s, which justifies the notation. Since $\mathcal{L}(H)$ is an Ulrich line bundle we have $\chi(\mathcal{L}) = \chi(\mathcal{L}(-H)) = 0$ which gives $L^2 = -4$ and $LH = g - 1$, where $L = c_1(\mathcal{L})$. Similar constraints hold for $H - L$. In particular, L and $H - L$ have the same degree with respect to H, hence $\mathrm{h}^0(\mathcal{O}_X(2L - H)) \le 1$, with equality being attained if and only if $L \equiv H - L$. Likewise, $\mathrm{h}^2(\mathcal{O}_X(2L - H)) = \mathrm{h}^0(\mathcal{O}_X(H - 2L)) \le 1$. Now we observe the following bound:

$$\begin{aligned}
\mathrm{ext}^1_X(\mathcal{L}^*(H), \mathcal{L}) &= \mathrm{h}^1(\mathcal{O}_X(2L - H)) \\
&= \mathrm{h}^0(\mathcal{O}_X(2L - H)) + \mathrm{h}^2(\mathcal{O}_X(2L - H)) - \chi(\mathcal{O}_X(2L - H)) \\
&\le 2 - \chi(\mathcal{O}_X(2L - H)) \\
&= g + 7,
\end{aligned}$$

the last equation being obtained by Riemann–Roch after plugging $L^2 = -4$ and $HL = g-1$. In view of the rigidity of $H - L$ and L, the family of sheaves appearing as an extension (5) is parametrized by $\mathbb{P}(\mathrm{Ext}^1_X(\mathcal{L}^*(H), \mathcal{L}))$ and hence has dimension at most $g+6$. So this family cannot dominate the $(2g+8)$-dimensional family S_0, a contradiction.

It follows from Theorem 1 that X is strictly Ulrich wild in the sense of [Faenzi and Pons-Llopis 2015]. The next result refines this fact in terms of moduli spaces. It was proved when $\mathrm{Pic}(X)$ is generated by H in [Aprodu et al. 2017, Theorem 2.7]. A modification of that argument allows to prove the result in general.

Theorem 7. *Let X be a $K3$ surface and H be a very ample line bundle on X. Then, for any positive integer r, the moduli space $\mathrm{M}_X(2r, rH, -2r)$ is of dimension $2(r^2(g+3)+1)$. Given a general sheaf $\mathcal{F}$ in this space, $\mathcal{F}(H)$ is a stable Ulrich bundle.*

Proof. Given a coherent sheaf $\mathcal{E}$ or rank $r > 0$ on X we write $\mathrm{P}(\mathcal{E}) \in \mathbb{Q}[t]$ for the Hilbert polynomial of $\mathcal{E}$ and $\mathrm{p}(\mathcal{E})$ for its reduced version, namely $\mathrm{P}(\mathcal{E}) = \chi(\mathcal{E}(tH))$ and $\mathrm{p}(\mathcal{E}) = \mathrm{P}(\mathcal{E})/r$. We put $\mathrm{p}_0 = (g-1)(t+1)t$ so that, if $\mathcal{E}$ is an Ulrich sheaf, then $\mathrm{p}(\mathcal{E}(-H)) = \mathrm{p}_0$. Note that, if $\mathcal{E}_1$ and $\mathcal{E}_2$ are nonisomorphic stable sheaves with $\mathrm{p}(\mathcal{E}_1) = \mathrm{p}(\mathcal{E}_2)$, then $\mathrm{Ext}^k_X(\mathcal{E}_i, \mathcal{E}_j) = 0$ for $k = 0, 2$ and $i \neq j$.

The proof goes by induction on r, the case $r = 1$ being given by Corollary 6. For $r \geq 1$, we select a stable bundle $\mathcal{E}_2$ in $\mathrm{M}_X(2r, rH, -2r)$ given by the induction hypothesis and a stable bundle $\mathcal{E}_1$ in $\mathrm{M}_X(2, H, -2)$, with $\mathcal{E}_i(H)$ Ulrich for $i = 1, 2$, taking care that $\mathcal{E}_1$ is not isomorphic to $\mathcal{E}_2$ for $r = 1$. This is of course possible since $\dim(\mathrm{M}_X(2, H, -2)) > 0$. This way we have:

$$\mathrm{Ext}^k_X(\mathcal{E}_i, \mathcal{E}_j) = 0, \qquad \text{for } k = 0, 2 \text{ and } i \neq j, \tag{6}$$

$$\mathrm{ext}^1_X(\mathcal{E}_i, \mathcal{E}_j) = 2r(g+3) \qquad \text{for } i \neq j. \tag{7}$$

Note that, for any choice of $\zeta \in \mathbb{P}(\mathrm{Ext}^1_X(\mathcal{E}_2, \mathcal{E}_1))$, the sheaf $\mathcal{E}^\zeta$ fitting as middle term of the associated extension is a locally free semistable sheaf, with $\mathcal{E}^\zeta(H)$ (as extension of sheaves having these properties). By direct computation, we see that it lies $\mathrm{M}_X(2(r+1), (r+1)H, -2(r+1))$. Of course this sheaf is not stable, as $\mathcal{E}_1$ is a subsheaf of $\mathcal{E}^\zeta$ with quotient $\mathcal{E}_2$ and the reduced Hilbert polynomial of all these sheaves is p_0. However, it follows by [Faenzi and Pons-Llopis 2015, Theorem A, ii)] that $\mathcal{E}^\zeta$ is simple, as the representation of the associated Kronecker consists of a single nonzero map of one-dimensional vector spaces, and as such it is simple. Alternatively one may apply [Pons-Llopis and Tonini 2009, Proposition 5.3].

We record the defining sequence:

$$0 \to \mathcal{E}_1 \to \mathcal{E}^\zeta \to \mathcal{E}_2 \to 0. \tag{8}$$

In the same spirit as in Lemma 5, we take a deformation of $\mathcal{E}^\zeta$ in the space of simple sheaves, which is unobstructed of dimension $2((r+1)^2(g+3)+1)$ at $\mathcal{E}^\zeta$. We consider thus an integral quasiprojective variety S as base of an S-flat family of simple sheaves $\mathcal{F}_s$ with $\mathcal{F}_s(H)$ Ulrich for all s and $\mathcal{F}_{s_0} \simeq \mathcal{E}^\zeta$

for some $s_0 \in S$, the base S being locally isomorphic to the moduli space of simple sheaves around the point s_0. We may assume that $\mathcal{F}_s$ is locally free for all $s \in S$.

Claim 8. *There is an open dense subset* S_0 *of* S *such that, for any stable sheaf* $\mathcal{K}$ *with* $\mathrm{rk}(\mathcal{K}) < 2(r+1)$, $\mathrm{rk}(\mathcal{K}) \neq 2$ *and* $\mathrm{p}(\mathcal{K}) = \mathrm{p}_0$, *we have* $\mathrm{Hom}_X(\mathcal{K}, \mathcal{F}_s) = 0$, *for all* $s \in S_0$.

Proof of the claim. Clearly it suffices to find such open subset for a fixed rank u of $\mathcal{K}$ and take the intersection of the corresponding open subsets for all $u < 2(r+1)$, $u \neq 2$.

So let N be the moduli space of stable sheaves $\mathcal{E}$ on X with Hilbert polynomial $\mathrm{P}(\mathcal{E}) = u\mathrm{p}_0$. Let $\mathcal{U}$ be a quasiuniversal family over $X \times \mathrm{N}$ [Huybrechts and Lehn 1997, Proposition 4.6.2] and denote by σ and π the projection maps $X \times \mathrm{N} \to \mathrm{N}$ and $X \times \mathrm{N} \to X$, respectively.

For $y \in \mathrm{N}$ let $\mathcal{U}_y$ be the corresponding sheaf over X. We observe that, applying $\mathrm{Hom}_X(\mathcal{U}_y, -)$ to (8), using the definition of N and ζ and the fact that the $\mathcal{E}_i$'s are stable with $\mathrm{p}(\mathcal{E}_i) = \mathrm{p}(\mathcal{U}_y)$ we get $\mathrm{Hom}_X(\mathcal{U}_y, \mathcal{E}^\zeta) = 0$. Indeed, the only case to check is for $u = 2r$ when y corresponds to the sheaf $\mathcal{E}_2$, but $\mathrm{Hom}_X(\mathcal{E}_2, \mathcal{E}^\zeta) = 0$, for otherwise by stability of $\mathcal{E}_2$ the exact sequence (8) would split, contradicting our assumption on ζ.

Then, Serre duality gives, for all $y \in \mathrm{N}$,

$$\mathrm{H}^2((\mathcal{E}^\zeta)^* \otimes \mathcal{U}_y) \simeq \mathrm{Ext}^2_X(\mathcal{E}^\zeta, \mathcal{U}_y) = 0. \tag{9}$$

Now consider $X \times \mathrm{N} \times S$, put τ for the projection $\mathrm{N} \times S \to S$ and denote by $\bar{\sigma}$, $\bar{\pi}$, $\bar{\tau}$ the projection maps from $X \times \mathrm{N} \times S$ onto $X \times S$, $\mathrm{N} \times S$ and $X \times \mathrm{N}$, respectively. Let $\mathcal{V} = \bar{\pi}^*(\mathcal{F}^*) \otimes \bar{\tau}^*(\mathcal{U})$. Since $\mathcal{V}$ is flat over the integral base $\mathrm{N} \times S$ and $\bar{\sigma}$ has relative dimension 2, base-change gives, for all $(y, s) \in \mathrm{N} \times S$

$$\boldsymbol{R}^2\bar{\sigma}_*(\mathcal{V})_{(y,s)} \simeq \mathrm{H}^2(\mathcal{F}_s^* \otimes \mathcal{U}_y). \tag{10}$$

Let W be the support of $\boldsymbol{R}^2\sigma_*(\mathcal{V})$, i.e., the closed subset of points $(y, s) \in \mathrm{N} \times S$ such that

$$\boldsymbol{R}^2\sigma_*(\mathcal{V})_{(y,s)} \neq 0.$$

By (9) and (10), we have $W \cap \mathrm{N} \times \{s_0\} = \varnothing$, i.e., s_0 does not lie in $\tau(W)$. Then there is an open neighborhood $S_0 \subset S$ of s_0 which is disjoint from $\tau(W)$. Again by (10), we get $\mathrm{H}^2(\mathcal{F}_s^* \otimes \mathcal{U}_y) = 0$ for all $(y, s) \in \mathrm{N} \times S_0$, which proves the claim. □

Let us now conclude the proof of the theorem. In view of the claim, we have two alternatives for s generic in S_0: either $\mathrm{Hom}(\mathcal{K}, \mathcal{F}_s) = 0$ for any stable sheaf $\mathcal{K}$ with $\mathrm{rk}(\mathcal{K}) < 2(r+1)$ and $\mathrm{p}(\mathcal{K}) = \mathrm{p}_0$ or otherwise this happens for all such $\mathcal{K}$ except for $\mathrm{rk}(\mathcal{K}) = 2$ and there actually exists a stable $\mathcal{K}$ in N such that $\mathrm{Hom}(\mathcal{K}, \mathcal{F}_s) \neq 0$.

In the first alternative $\mathcal{F}_s$ is stable, so we assume that the second one takes place and look for a contradiction. We go back to Claim 8 and carry out the same argument for $u = 2$, with y_0 being the point corresponding to $\mathcal{E}_1$. Observe that $\mathcal{K}$ must lie in $\mathrm{M}_X(2, H, -2)$ as the proof of Claim 8 applies verbatim on any other component of N.

We note that $W \cap \mathrm{N} \times \{s_0\} = \{(y_0, s_0)\}$, as clearly $\mathrm{Hom}_X(\mathcal{K}, \mathcal{E}^\zeta) = 0$ for all $\mathcal{K}$ in $\mathrm{N} \setminus \{y_0\}$. So W is properly contained in $\mathrm{N} \times S$. Moreover, we easily have $\hom_X(\mathcal{E}_1, \mathcal{E}^\zeta) = 1$. Recall by construction of the quasiuniversal family that there is u_0 such that $\mathrm{rk}(\mathcal{U}) = 2u_0$ and that, for $y \in \mathrm{N}$, the sheaf $\mathcal{U}_y$ is a direct sum of u_0 copies of the stable sheaf of rank 2 in $\mathrm{M}_X(2, H, -2)$ corresponding to y. Therefore, the sheaf $\boldsymbol{R}^2\bar{\sigma}_*(\mathcal{V})_{(y,s)}$ has rank at least u_0 at any $(y, s) \in W$, and rank precisely u_0 at (y_0, s_0). So there is an open dense subset W_0 of W where $\boldsymbol{R}^2\bar{\sigma}_*(\mathcal{V})$ is free of rank u_0. For any $(y, s) \in W_0$, the stable sheaf $\mathcal{K}$ corresponding to y satisfies $\hom_X(\mathcal{K}, \mathcal{F}_s) = 1$; up to proportionality we have thus a unique nonzero map $\eta_{y,s} : \mathcal{K} \to \mathcal{F}_s$. Stability easily implies that $\eta_{y,s}$ is injective, so there is an exact sequence

$$0 \to \mathcal{K} \to \mathcal{F}_s \to \mathcal{K}' \to 0,$$

for a well-defined sheaf $\mathcal{K}' = \mathrm{coker}(\eta_{y,s})$, for all $(y, s) \in W_0$.

For $s = s_0$ the sheaf $\mathcal{K}'$ is just $\mathcal{E}_2$ so, by openness of stability, up to shrinking W_0 we may assume that $\mathcal{K}'$ is stable for all $(y, s) \in W_0$. Note that $\mathcal{K}'$ lies in $\mathrm{M}(2r, rH, -2r)$.

Under our assumption, such sequence should exist for any s in an open neighborhood of s_0. Then the family of sheaves $\mathcal{F}$ should be dominated by the family of extensions of $\mathcal{K}$ by $\mathcal{K}'$ as s varies around s_0. We see that the dimension of this family of extensions is

$$\dim(\mathrm{M}_X(2, H, -2)) + \dim(\mathrm{M}_X(2r, rH, -2r)) + \dim(\mathbb{P}\,\mathrm{Ext}^1_X(\mathcal{K}', \mathcal{K})),$$

which equals $2(r(r+1)+1)(g+3)+3$, as it follows by formulas (6) and (7) applied to $\mathcal{K}$ and $\mathcal{K}'$ instead of $\mathcal{E}_1$ and $\mathcal{E}_2$. On the other hand, the dimension of S is $2((r+1)^2(g+3)+1)$. The difference of these dimensions is $2r(g+3)-1$ and since this is always positive for $r \geq 1$, $g \geq 3$, we get that the family of simple sheaves appearing as extensions cannot be dense in S_0. This contradiction concludes the proof. ☐

The previous result is in some sense optimal as general K3 surfaces do not support Ulrich bundles of odd rank [Aprodu et al. 2017, Corollary 2.2].

Remark. An argument similar to the one of Theorem 1 has been used to construct ACM and Ulrich bundles on Fano threefolds of index 1. Indeed, it follows from the main result of [Brambilla and Faenzi 2011] that any smooth Fano threefold of Picard number 1 and index 1, containing a line L with normal bundle $\mathcal{O}_L \oplus \mathcal{O}_L(-1)$ (such a threefold was called “ordinary” in that paper) admits an Ulrich bundle of rank 2. Ulrich sheaves of rank 2 are precisely ACM sheaves $\mathcal{E}$ with $c_1(\mathcal{E}(-H)) = H$ and $c_2(\mathcal{E}(-H)) = (g+3)L$, where $L \subset X$ is a line. We do not know if the same result holds for nonordinary threefolds.

Remark. Theorem 1 implies for instance that any integral quartic surface supports an Ulrich bundle of rank 2. If X is not integral, then X must the union of (possibly multiple) surfaces of degree ≥ 3. For each component it is possible to find a rank-2 Ulrich bundle, we refer to [Faenzi and Pons-Llopis 2015, Lemma 7.2] for the slightly delicate case of singular cubic surfaces. This yields existence of an Ulrich sheaf of rank 2 on an arbitrary quartic surface.

However the resulting sheaf will fail to be locally free over the intersection of the components. Finding locally free Ulrich sheaves of rank 2 seems more tricky when X is not irreducible and might be impossible when X is not reduced. To justify this let us mention that, for instance if X the union of two distinct double planes, the rank of any locally free Ulrich sheaf on X must be a multiple of 4 by [Ballico et al. 2019, Proposition 4.14].

Acknowledgements

I would like to thank M. Aprodu, G. Casnati, A. Perego, J. Pons-Llopis and P. Stellari for useful discussions. I am grateful to the referee for useful remarks.

References

[Altman and Kleiman 1980] A. B. Altman and S. L. Kleiman, "Compactifying the Picard scheme", *Adv. Math.* **35**:1 (1980), 50–112. MR Zbl

[Aprodu et al. 2017] M. Aprodu, G. Farkas, and A. Ortega, "Minimal resolutions, Chow forms and Ulrich bundles on K3 surfaces", *J. Reine Angew. Math.* **730** (2017), 225–249. MR Zbl

[Artamkin 1990] I. V. Artamkin, "Deformation of torsion-free sheaves on an algebraic surface", *Izv. Akad. Nauk SSSR Ser. Mat.* **54**:3 (1990), 435–468. In Russian; translated in *Math. USSR-Izv.* **36**:3 (1991), 449–485. MR Zbl

[Ballico et al. 2019] E. Ballico, S. Huh, F. Malaspina, and J. Pons-Llopis, "ACM sheaves on the double plane", *Trans. Amer. Math. Soc.* **372**:3 (2019), 1783–1816. MR Zbl

[Beauville 2016] A. Beauville, "Ulrich bundles on abelian surfaces", *Proc. Amer. Math. Soc.* **144**:11 (2016), 4609–4611. MR Zbl

[Beauville 2018] A. Beauville, "An introduction to Ulrich bundles", *Eur. J. Math.* **4**:1 (2018), 26–36. MR Zbl

[Brambilla and Faenzi 2011] M. C. Brambilla and D. Faenzi, "Moduli spaces of rank-2 ACM bundles on prime Fano threefolds", *Michigan Math. J.* **60**:1 (2011), 113–148. MR Zbl

[Casnati 2017] G. Casnati, "Special Ulrich bundles on non-special surfaces with $p_g = q = 0$", *Int. J. Math.* **28**:8 (2017), art. id. 1750061. MR Zbl

[Coskun 2017] E. Coskun, "A survey of Ulrich bundles", pp. 85–106 in *Analytic and algebraic geometry* (Mumbai/Hyderabad, 2015), edited by A. Aryasomayajula et al., Hindustan Book Agency, New Delhi, 2017. MR Zbl

[Coskun et al. 2012] E. Coskun, R. S. Kulkarni, and Y. Mustopa, "Pfaffian quartic surfaces and representations of Clifford algebras", *Doc. Math.* **17** (2012), 1003–1028. MR Zbl

[Eisenbud et al. 2003] D. Eisenbud, F.-O. Schreyer, and J. Weyman, "Resultants and Chow forms via exterior syzygies", *J. Amer. Math. Soc.* **16**:3 (2003), 537–579. MR Zbl

[Faenzi and Pons-Llopis 2015] D. Faenzi and J. Pons-Llopis, "The Cohen–Macaulay representation type of arithmetically Cohen–Macaulay varieties", preprint, 2015. arXiv

[Huybrechts and Lehn 1997] D. Huybrechts and M. Lehn, *The geometry of moduli spaces of sheaves*, Aspects of Math. **E31**, Vieweg & Sohn, Braunschweig, Germany, 1997. MR Zbl

[Kollár 2008] J. Kollár, "Hulls and husks", preprint, 2008. arXiv

[Mukai 1984] S. Mukai, "Symplectic structure of the moduli space of sheaves on an abelian or K3 surface", *Invent. Math.* **77**:1 (1984), 101–116. MR Zbl

[Pons-Llopis and Tonini 2009] J. Pons-Llopis and F. Tonini, "ACM bundles on del Pezzo surfaces", *Matematiche* (*Catania*) **64**:2 (2009), 177–211. MR Zbl

[Qin 1993] Z. Qin, "Moduli of simple rank-2 sheaves on K3-surfaces", *Manuscripta Math.* **79**:3-4 (1993), 253–265. MR Zbl

[Rosenlicht 1952] M. Rosenlicht, "Equivalence relations on algebraic curves", *Ann. of Math.* (2) **56** (1952), 169–191. MR Zbl

[Saint-Donat 1974] B. Saint-Donat, "Projective models of K-3 surfaces", *Amer. J. Math.* **96** (1974), 602–639. MR Zbl

[Schreyer and Eisenbud 2010] F.-O. Schreyer and D. Eisenbud, "Betti numbers of syzygies and cohomology of coherent sheaves", pp. 586–602 in *Proceedings of the International Congress of Mathematicians, II* (Hyderabad, 2010), edited by R. Bhatia et al., Hindustan Book Agency, New Delhi, 2010. MR Zbl

Communicated by Gavril Farkas
Received 2018-10-07 Revised 2019-02-07 Accepted 2019-03-11

daniele.faenzi@u-bourgogne.fr *Institut de Mathématiques de Bourgogne, UMR CNRS 5584, Université de Bourgogne, Dijon, France*

ALGEBRA AND NUMBER THEORY 13:6 (2019)
dx.doi.org/10.2140/ant.2019.13.1455

Unlikely intersections in semiabelian surfaces

Daniel Bertrand and Harry Schmidt

We consider a family, depending on a parameter, of multiplicative extensions of an elliptic curve with complex multiplications. They form a 3-dimensional variety G which admits a dense set of special curves, known as Ribet curves, which strictly contains the torsion curves. We show that an irreducible curve W in G meets this set Zariski-densely only if W lies in a fiber of the family or is a translate of a Ribet curve by a multiplicative section. We further deduce from this result a proof of the Zilber–Pink conjecture (over number fields) for the mixed Shimura variety attached to the threefold G, when the parameter space is the universal one.

1. Introduction

1.1. ***Statement of the results and plan of the proofs.*** Let $E_0/\mathbb{Q}^{\mathrm{alg}}$ be an elliptic curve with complex multiplications. On any extension G_0 of E_0 by $\mathbb{G}_m$ defined over $\mathbb{Q}^{\mathrm{alg}}$, there exists a particular subgroup Γ_0 of $G_0(\mathbb{Q}^{\mathrm{alg}})$, whose elements are called Ribet points. We refer to Section 1.2 below for their precise definition, but point out right now that Γ_0 contains the torsion subgroup G_0^{tor} of $G_0(\mathbb{Q}^{\mathrm{alg}})$. In fact $\Gamma_0 = G_0^{\mathrm{tor}}$ if the extension G_0 is isosplit, while Γ_0 has rank 1 otherwise.

Let further $X/\mathbb{Q}^{\mathrm{alg}}$ be a smooth irreducible algebraic curve and let G/X be an X-extension of $E_{0/X}$ by $\mathbb{G}_{m/X}$. Let q be the section of $\hat{E}_{0/X} \to X$ representing the isomorphism class of the extension G/X. We identify q with its image in $E_0(X)$ under the standard polarization $\hat{E}_0 \simeq E_0$, and write $G \simeq G_q$. Given a section s of G/X, we denote by $p = \pi \circ s \in E_0(X)$ its composition with the projection $\pi : G \to E_0 \times X$.

Let $\delta \neq 0$ be a purely imaginary complex multiplication of E_0, and let $\xi \in X(\mathbb{Q}^{\mathrm{alg}})$. A first property of Ribet points is that if $s(\xi)$ is a Ribet point of its fiber $G_\xi \simeq G_{q(\xi)}$, then its projection $p(\xi)$ to E_0 and the point $\delta q(\xi)$ are linearly dependent over $\mathbb{Z}$. Usually, this condition alone will be satisfied by infinitely

MSC2010: primary 14K15; secondary 11G15, 11G50, 11U09.
Keywords: semiabelian varieties, complex multiplication, Zilber–Pink conjecture, mixed Shimura varieties, heights, o-minimality, Ribet sections.

many ξ's. But asking that $s(\xi)$ be a Ribet point in the fiber of $G_\xi \to E_0$ above $p(\xi)$ brings a second condition, unlikely to be satisfied infinitely often. And indeed, we prove in this paper:

Theorem 1. *Let $G \simeq G_q$ be a nonconstant (hence nonisosplit) extension of $E_{0/X}$ by $\mathbb{G}_{m/X}$, and let s be a section of $G \to X$, all defined over $\mathbb{Q}^{\mathrm{alg}}$. Assume that the set*

$$\Xi = \Xi_s := \{\xi \in X(\mathbb{Q}^{\mathrm{alg}}) \mid s(\xi) \textit{ is a Ribet point of its fiber } G_\xi \simeq G_{q(\xi)}\}$$

is infinite. Then, the sections p and q are linearly dependent over $\mathrm{End}(E_0)$.

Referring again to Section 1.2 for the definition of the Ribet sections of G/X (which in view of the hypothesis on G, also form a group Γ of rank 1, containing the torsion sections), we deduce the following (actually equivalent) version of Theorem 1:

Theorem 2. *Assume that the hypotheses of Theorem 1 on the extension G, the section s and the set Ξ are satisfied. Then, there exists a nonconstant or trivial section s' in $\mathbb{G}_m(X)$ such that $s - s'$ is a Ribet section of G/X.*

The conclusion of Theorem 2 is best possible. Indeed, let s' be such a section in $\mathbb{G}_m(X)$ and let s'' be a Ribet section. Then, $s''(\xi)$ is a Ribet point of G_ξ for any $\xi \in X$, while $s'(\xi)$ lies in $\mathbb{G}_m^{\mathrm{tor}}$ infinitely often. The set Ξ_s attached to $s = s' + s''$ is therefore infinite.

As a corollary to Theorem 1, we consider the case when the curve $X = \hat{E}_0 \simeq \mathrm{Ext}(E_0, \mathbb{G}_m)$ is the parameter space of the universal extension $\mathcal{P}_0$ of E_0 by $\mathbb{G}_m$. This extension, which identifies with the Poincaré biextension of $E_0 \times \hat{E}_0$ by $\mathbb{G}_m$, is naturally endowed with the structure of a mixed Shimura variety, for which we prove:

Theorem 3. *Let $W/\mathbb{Q}^{\mathrm{alg}}$ be an irreducible algebraic curve in $\mathcal{P}_0$. Assume that W contains infinitely many points lying on special curves of the mixed Shimura variety $\mathcal{P}_0$. Then, W is contained in a special surface of $\mathcal{P}_0$.*

Combined with Gao's work on the André–Oort conjecture, this readily implies the following conclusion, which answers a question of J. Pila.

Theorem 4. *The mixed Shimura variety $\mathcal{P}_0$ satisfies the Zilber–Pink conjecture over number fields.*

See Section 5 below for the statement of this conjecture, and for the deduction of Theorems 3 and 4 from Theorem 1.

The proof of Theorem 1 will distinguish three cases. In the first one, we establish the following weaker version, where the conclusion is replaced by a "weakly special" one. Denote by $E_0(\mathbb{Q}^{\mathrm{alg}}) \subset E_0(X)$ the group of constant sections of $E_{0/X}$.

Theorem 1.w. *Same hypotheses as in Theorem 1. Then, the sections p and q are linearly dependent over* $\mathrm{End}(E_0)$ *modulo* $E_0(\mathbb{Q}^{\mathrm{alg}})$.

The proof of Theorem 1.w (see Section 2) follows the o-minimal strategy of Pila–Zannier and Masser–Zannier, starting with the observation that if its conclusion does not hold, then the points ξ of Ξ have bounded height.

In the remaining cases, we suppose that p and q are linearly dependent over $\mathrm{End}(E_0)$ modulo $E_0(\mathbb{Q}^{\mathrm{alg}})$. In the second one (see Section 3), we assume that they are linearly dependent over $\mathbb{Z}$ modulo $E_0(\mathbb{Q}^{\mathrm{alg}})$, but that p is not (i.e., p is not constant). Here again, we use the o-minimal strategy, but a new argument is required to check bounded height.

In the last case (see Section 4), we reduce a weakly special relation over $\mathrm{End}(E_0)$ to one over $\mathbb{Z}$, and therefore to a constant section p. We finally show that p must be torsion, thanks to a duality argument which turns the problem into a special case of the Mordell–Lang theorem (recalled in Section 1.3(v) below) for a constant semiabelian variety attached not to q, but to p.

1.2. *Ribet sections and points.* Let $\mathcal{X}/\mathbb{Q}^{\mathrm{alg}}$ be a smooth irreducible variety, let A be an abelian scheme over $\mathcal{X}$, let $q \in \hat{A}(\mathcal{X})$ be a section of the dual abelian scheme $\hat{A}/\mathcal{X} \simeq \mathrm{Ext}_{\mathcal{X}}(A, \mathbb{G}_m)$, and let $G = G_q$ be the corresponding $\mathcal{X}$-extension of A by $\mathbb{G}_{m/\mathcal{X}}$, obtained by removing its zero section from the line bundle defined by q. We point out that G_q is an isosplit extension (i.e., isogenous to the product $\mathbb{G}_m \times A$) if and only if q is a torsion section. When $A/\mathcal{X}$ is a constant group scheme, G_q is a constant group scheme if and only if q is a constant section (for instance a torsion one).

Let $\mathcal{P}$ be the Poincaré biextension of $A \times_{\mathcal{X}} \hat{A}$ by $\mathbb{G}_m$. For any $\varphi \in \mathrm{Hom}_{\mathcal{X}}(\hat{A}, A)$, with transpose $\hat{\varphi}$, there is a canonical isomorphism $\sigma_{\varphi,q} : \mathcal{P}((\varphi - \hat{\varphi})(q), q) \simeq \mathbb{G}_{m/\mathcal{X}}$ of $\mathbb{G}_m$-torsors over $\mathcal{X}$ (see [Chambert-Loir 1999, Proposition 6.3], whose description of $\sigma_{\varphi,q}$ works over an arbitrary base scheme [Bertrand and Edixhoven 2019, Proposition 3.1]). We define the *basic Ribet section* associated to φ as the section $s_{\varphi,q} = \sigma^*_{\varphi,q}(1_{\mathcal{X}})$ of the semiabelian scheme $G = G_q = (\mathrm{id}_A, q)^*\mathcal{P} = \mathcal{P}_{|A\times q}$ over $\mathcal{X}$. We say "point" instead of "section" if $\mathcal{X}$ is a point, and drop the index q when the context is clear.

The Ribet section $s_\varphi \in G(\mathcal{X})$ depends additively on φ, and in fact only on $\varphi - \hat{\varphi}$ [Jacquinot and Ribet 1987, Proposition 4.2; Bertrand and Edixhoven 2019, Formula 3.1.2]. Its projection under $\pi : G \to A$ is the section

$$p_\varphi := \pi \circ q_\varphi = (\varphi - \hat{\varphi}) \circ q \in A(\mathcal{X}).$$

So, when φ varies, the basic Ribet sections form a finitely generated subgroup of $G(\mathcal{X})$, of rank r_q at most equal to the rank of the $\mathbb{Z}$-module $\mathcal{E} = \{\varphi - \hat{\varphi}, \varphi \in \mathrm{Hom}_{\mathcal{X}}(\hat{A}, A)\}$, and equal to it when q is sufficiently general. On the other hand, $r_q = 0$ if q is a torsion section. Indeed, although their dependence in q is *not linear*, the Ribet sections s_φ satisfy the following "lifting property" (for (i) $\Rightarrow$ (ii), see [Bertrand 2011, §1], [Bertrand et al. 2016, Theorem 3(i)] in the case of points, and [Bertrand and Edixhoven 2019, Proposition 3.3] in general).

Lemma 1. *Let $\varphi \in \mathrm{Hom}_{\mathcal{X}}(\hat{A}, A)$, let $q \in \hat{A}(\mathcal{X})$ and consider the conditions*:

(i) *q is a torsion section.*

(ii) *s_φ is a torsion section.*

(iii) *p_φ is a torsion section.*

Then, (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii), *and if $\varphi - \hat{\varphi}$ is an isogeny, the three conditions are equivalent.*

More generally, let s be a local section of $G \to \mathcal{X}$ (for the étale topology). We say that s is a *Ribet section* of $G/\mathcal{X}$ if there exists a positive integer n satisfying $n.s = s_\varphi$ for some φ, with multiplication by n in the sense of the group scheme $G/\mathcal{X}$. The projection p of s to A satisfies $np = (\varphi - \hat{\varphi}) \circ q$. All (local) torsion sections of $G/\mathcal{X}$ now appear as such Ribet sections, and Lemma 1 extends to this more general setting. Viewed as points above the generic point η of $\mathcal{X}$, with $K = \mathbb{Q}^{\mathrm{alg}}(\mathcal{X}_\eta)$, the Ribet sections form a subgroup Γ of the group $G_\eta(K^{\mathrm{alg}})$, of same rank r_q as above.

The construction of Ribet sections commutes with any base change. For instance, given a basic Ribet section $s_{\varphi,q}$ of $G/\mathcal{X}$, and a point ξ in $\mathcal{X}(\mathbb{Q}^{\mathrm{alg}})$, $s_{\varphi,q}(\xi) = s_{\varphi_\xi, q(\xi)}$ is the basic Ribet point of the fiber G_ξ attached to the specialization φ_ξ of φ at ξ. Conversely, let s^ξ be a Ribet point of $G_\xi(\mathbb{Q}^{\mathrm{alg}})$. By definition, there exist $n_\xi \in \mathbb{Z}_{>0}$ and $\varphi_\xi \in \mathrm{Hom}(\hat{A}_\xi, A_\xi)$ such that $n_\xi s^\xi = s_{\varphi_\xi, q(\xi)}$. *Assume further that φ_ξ extends to an element $\varphi \in$* $\mathrm{Hom}(\hat{A}, A)$ (which occurs automatically if $A/\mathcal{X}$ is a constant abelian scheme as in Section 1.1). Then, $s_{\varphi_\xi, q(\xi)} = s_{\varphi,q}(\xi)$, and there exists a local section s of $G/\mathcal{X}$ such that $n_\xi.s = s_\varphi$, whose image in G contains s^ξ. So, the Ribet point s^ξ extends locally to a Ribet section of $G/\mathcal{X}$.

Let us now return to the situation of Section 1.1, where $A = E_0 \times \mathcal{X}$, for a CM elliptic curve E_0, and $\mathcal{X}$ is either the curve X or a point ξ on X. Then, the $\mathbb{Z}$-module $\mathcal{E}$ above identifies with

$$\mathcal{E} = \{\varphi - \bar{\phi} \mid \varphi \in \mathrm{End}(E_0)\} = \mathbb{Z}\delta,$$

where $\delta = \alpha - \bar{\alpha} \neq 0$ is a purely imaginary quadratic number, which will be fixed from now on. Consequently, for any $q \in E_0(X)$, the group of basic Ribet sections of $G = G_q$ is cyclic, generated by the section

$$s^R := s_{\alpha,q} \in G(X), \quad \text{with } p^R := \pi \circ s^R = \delta q \in E_0(X).$$

Viewed at the generic point η of X, *the Ribet sections of G/X then form the divisible hull Γ of the group $\mathbb{Z}.s^R(\eta)$ in $G_\eta(K^{\mathrm{alg}})$*. Furthermore, for any $\xi \in X(\mathbb{Q}^{\mathrm{alg}})$, the value $s^R(\xi) = s_{\alpha,q(\xi)}$ of s^R at ξ generates the group of basic Ribet sections of $G_\xi = G_{q(\xi)}$, and the Ribet points of G_ξ form the divisible hull

$$\Gamma_\xi = \{s^\xi \in G_\xi(\mathbb{Q}^{\mathrm{alg}}) \mid \exists (n,m) \in \mathbb{Z}^2, n \neq 0, ns^\xi = ms^R(\xi)\} \supset G_\xi^{\mathrm{tor}}$$

of $\mathbb{Z}.s^R(\xi)$ in $G_\xi(\mathbb{Q}^{\mathrm{alg}})$.

Under the assumptions of Section 1.1, the section q is not constant, hence not torsion, while δ is an isogeny, so s^R is not torsion by Lemma 1, and the rank r_q of Γ is equal to 1. On the other hand, by Lemma 1 (now at the level of points), given a point $\xi \in X(\mathbb{Q}^{\mathrm{alg}})$,

$$q(\xi) \in E_0^{\mathrm{tor}} \Leftrightarrow s^R(\xi) \in G_\xi^{\mathrm{tor}} \Leftrightarrow \Gamma_\xi = G_\xi^{\mathrm{tor}},$$

and this occurs for *infinitely many* ξ's since q is not constant [Bertrand 2011, Theorem 1]. Otherwise, Γ_ξ has rank 1, but for $s(\xi) \in \Gamma_\xi$, we still have $s(\xi) \in G_\xi^{\mathrm{tor}} \Leftrightarrow p(\xi) \in E_0^{\mathrm{tor}}$.

In view of these descriptions of the groups Γ and Γ_ξ, our work can be interpreted as a particular case of the study of unlikely intersections within an isogeny class [Gao 2017a], or of a relative version of the Mordell–Lang problem (compare with Section 1.3(v) below).

1.3. ***The context.*** We here put the results of Section 1.1 in perspective with other statements of unlikely intersections. Two sets

$$\Xi^{\mathrm{tor}} \subset \Xi \subset \Xi^{\ell d}$$

related to the section $s \in G(X)$ naturally appear in the process.

(i) Theorem 1 gives a positive answer to the "Question 2" raised in [Bertrand 2013, §5], while a positive answer to its "Question 1" was recently obtained by Barroero [2017]. However, the applications to Pink's conjecture given in [Bertrand 2013] require clarification, because of their ambiguous use of Hecke orbits. We bypass this problem for the mixed Shimura variety $\mathcal{P}_0$ studied in Section 5, by describing all its possible special curves. Theorem 3 will then follow from Theorem 1, along the method of [Bertrand 2013].

(ii) Contrary to the convention of [Bertrand et al. 2016], the torsion points are here viewed as particular cases of Ribet points. Therefore, Theorem 2 implies the restriction to the case of our semiabelian scheme G/X of the main theorem of [Bertrand et al. 2016], which concerns the subset

$$\Xi^{\mathrm{tor}} = \Xi_s^{\mathrm{tor}} := \{\xi \in X(\mathbb{Q}^{\mathrm{alg}}),\, s(\xi) \text{ is a torsion point of its fiber } G_\xi\}$$

of Ξ, and asserts the following statement.

Lemma 2. *Let G/X and s be as in Theorem 1, and assume moreover that the subset Ξ^{tor} of Ξ is infinite. Then s is a Ribet section or a torsion translate of a nonconstant section in $\mathbb{G}_m(X)$.*

For $\xi \in \Xi^{\mathrm{tor}}$, $p(\xi)$ too is torsion, so (by the Manin–Mumford theorem [Hindry 1988] for the image of (p, s') in $E_0 \times \mathbb{G}_m$), the conclusion of Theorem 2 can be sharpened to the same statement.

Let $\Xi_{s^R}^{\mathrm{tor}}$ be the set attached to the Ribet section s^R, defined similarly as Ξ_s^{tor}. We pointed out at the end of Section 1.2 that $\Xi_{s^R}^{\mathrm{tor}}$ is infinite. Therefore, Lemma 2 too is best possible.

(iii) In relation with the two sections s, s^R of G/X, consider the set

$$\Xi^{\ell d} = \Xi_{s,s^R}^{\ell d} := \{\xi \in X(\mathbb{Q}^{\mathrm{alg}}) \mid s(\xi) \text{ and } s^R(\xi) \text{ are linearly dependent over } \mathbb{Z}\}.$$

For ξ in this set, either $s(\xi)$ lies in the divisible hull Γ_ξ of $\mathbb{Z}.s^R(\xi)$, or $s^R(\xi)$ is a torsion point. So $\Xi^{\ell d}$ is the (not necessarily disjoint) union of Ξ and $\Xi_{s^R}^{\mathrm{tor}}$ and in particular, is always infinite. More generally, given two sections s, s' in $G(X)$, the similarly defined set $\Xi_{s,s'}^{\ell d}$ will be infinite as soon as the group generated by s and s' in $G(X)$ contains a nontorsion Ribet section. So, in contrast with the case of abelian schemes (see [Masser and Zannier 2015; Barroero and Capuano 2018]), the subgroup schemes of $G \times_X G$ do not suffice to control the finiteness of $\Xi_{s,s'}^{\ell d}$; as in [Bertrand and Edixhoven 2019], the special subvarieties of the corresponding mixed Shimura variety should also be taken into account.

(iv) Consider the curve $W = s(X)$ in G and define a Ribet curve as the image in G of a Ribet section. Theorem 2 then says that W is the translate of a Ribet curve by a section in $\mathbb{G}_m(X)$. Since any curve W in G dominating X can be viewed as the image of a section after a base extension, while any Ribet point of a fiber G_ξ locally extends to a Ribet section, this justifies the last but one sentence of the abstract.

(v) Assume that contrary to the hypothesis of Theorem 1, $G = G_0 \times X$ for some constant semiabelian surface $G_0/\mathbb{Q}^{\mathrm{alg}}$, and that s is not constant. Then, the projection W_0 of $W = s(X)$ to G_0 is a curve, which contains infinitely many points of the group Γ_0 of Ribet points of G_0. Since Γ_0 has finite rank (at most 1), the solution by Vojta and McQuillan [McQuillan 1995] of the *Mordell–Lang* conjecture for semiabelian varieties implies that s factors through a translate by a Ribet point of a strict connected algebraic subgroup of G_0. If the section q, here constant, is not torsion, the only such one is $\mathbb{G}_m$. So the conclusions of Theorems 1 and 2 still hold true in this case.

(vi) Same as in (v), but assume furthermore that q is a torsion section, say the trivial one, so $G_0 \simeq \mathbb{G}_m \times E_0$. Then, $s = (s', p)$ for some section $s' \in \mathbb{G}_m(X)$, while the group Γ_0 of Ribet points of G_0 coincides with G_0^{tor}. By Manin–Mumford, $\Xi = \Xi^{\mathrm{tor}}$ is then infinite if and only if s' is a torsion section, or p is a torsion section.

(vii) In this paper, we do not touch on the question of replacing $\mathbb{Q}^{\mathrm{alg}}$ by $\mathbb{C}$, or of applying Theorem 2 to generalized Pell equations as in [Masser and Zannier 2015; Barroero and Capuano 2018]. Nor do we study how effective our results can be made. Note that Lemma 2 above is made effective in the ongoing work [Jones and Schmidt $\geq$ 2019]. Due to the use of Pfaffian methods, in particular [Jones and Thomas 2018; Jones and Schmidt 2017], the bounds for the counting problem in [Jones and Schmidt $\geq$ 2019] are uniform and effective.

We take opportunity of these comments to show the following equivalence:

Theorem 1 $\Leftrightarrow$ *Theorem 2*. Theorem 2 clearly implies Theorem 1. Indeed, the sections s and $s'' = s - s'$ have the same projection p to E_0. Since s'' is a Ribet section, p and δq are linearly dependent over $\mathbb{Z}$, so p and q are linearly dependent over $\mathrm{End}(E_0)$.

Conversely, assume that the hypotheses and the conclusion of Theorem 1 hold true, and let $np - \rho q = 0$ be a nontrivial relation with $n \in \mathbb{Z}$, $\rho \in \mathrm{End}(E_0)$ not both 0 (equivalently, $n \neq 0$ since q is not a torsion section). Without loss of generality, we can assume that Ξ^{tor} is finite, otherwise Lemma 2 readily implies the conclusion of Theorem 2. For any $\xi \in \Xi$, $\delta q(\xi)$ and the projection $p(\xi)$ of the Ribet point $s(\xi)$ are linearly dependent over $\mathbb{Z}$, so there exist $n_\xi, m_\xi \in \mathbb{Z}$, not both zero, such that $n_\xi p(\xi) - m_\xi \delta q(\xi) = 0$, while the generic relation implies $np(\xi) - \rho q(\xi) = 0$. If these two relations are linearly independent over $\mathrm{End}(E_0)$, then $q(\xi)$, hence $s^R(\xi)$, hence $s(\xi)$, are torsion points and ξ lies in Ξ^{tor}. So, for infinitely many, hence at least one, ξ, these two relations must be linearly dependent over $\mathrm{End}(E_0)$, and in fact over $\mathbb{Z}$, since n does not vanish. This implies that ρ is a rational multiple of δ, and by their very construction, this in turn implies the existence of a Ribet section s'' projecting to p. So, $s' = s - s''$ factors through $\mathbb{G}_m$.

Finally, if s' is a constant section, it must be a torsion one since $s'(\xi)$ is a Ribet point of G_ξ projecting to 0 for one (any) $\xi \in \Xi$. In this case, s itself is a Ribet section, and otherwise s' is not constant, so the conclusion of Theorem 2 holds in all cases. □

2. Proof of Theorem 1.w

Recall the hypotheses of Theorem 1.w, as well as the notation $s^R, \Gamma_\xi, \dots$ of Section 1.2. So, $q \in E_0(X)$ is not constant, s is a section of $G = G_q \to X$ projecting to the section $\pi \circ s = p \in E_0(X)$, and the set $\Xi = \{\xi \in X(\mathbb{Q}^{\mathrm{alg}}), s(\xi) \in \Gamma_\xi\}$, concretely described as

$$\Xi = \{\xi \in X(\mathbb{Q}^{\mathrm{alg}}) \mid \exists (n, m) \in \mathbb{Z}^2,\ n \neq 0,\ ns(\xi) - ms^R(\xi) = 0\}$$

is infinite. *We assume that the sections p and q are linearly independent over* $\mathrm{End}(E_0)$ *modulo* $E_0(\mathbb{Q}^{\mathrm{alg}})$, and search for a contradiction.

We fix a number field k over which X and G, hence the sections q and s^R, as well as the section s, hence p, and the isogeny δ, are defined. We recall that the basic Ribet section s^R projects to E_0 on the section $p^R = \delta q$.

2.1. ***The o-minimal strategy.*** The proof of Theorem 1.w will be done in 5 steps. The third one is developed in Section 2.2. By a "constant" c, γ, we mean a positive real number which depends only on the data X, E_0, q, s and the number field k. The constants C may depend on further data introduced in the proof.

We point out that any finite set of points can without loss of generality be withdrawn from the curve X. To ease a technical point in the third step, we will for instance require that the sections p, q and $p + q \in E_0(X)$ never vanish on X. The complement is a finite set since q is not constant, p can be assumed to be so (constant p's are treated by a direct method in Section 4.2), and if $p + q$ is constant, we can make it nonconstant by replacing s by $2s$, so p by $2p$, without modifying the content of the theorems.

2.1.1. *Bounded heights of points.* Let h denote a height on $X(\mathbb{Q}^{\mathrm{alg}})$ attached to a divisor of degree 1 on the completed curve. Consider the set

$$\Xi^{\mathbb{Z}\ell d}_{p,\delta q} = \{\xi \in X(\mathbb{Q}^{\mathrm{alg}}) \mid p(\xi) \text{ and } \delta q(\xi) \text{ are linearly dependent over } \mathbb{Z}\}.$$

Since the projection $p(\xi) = \pi \circ s(\xi)$ of a Ribet point $s(\xi)$ lies in the divisible hull of the group $\mathbb{Z}.\delta q(\xi)$ in $E_0(\mathbb{Q}^{\mathrm{alg}})$, this set contains Ξ.

Lemma 3. *Let $p, q \in E_0(X)$ be linearly independent over* $\mathrm{End}(E_0)$ *modulo* $E_0(\mathbb{Q}^{\mathrm{alg}})$. *There exists a constant c_0 such that $h(\xi) \leq c_0$ for any $\xi \in \Xi^{\mathbb{Z}\ell d}_{p,\delta q}$, and in particular, for any $\xi \in \Xi$.*

Proof. In view of the hypothesis on p, q, bounded height on $\Xi^{\mathbb{Z}\ell d}_{p,\delta q}$ follows directly from [Viada 2003, Theorem 4] (and one can even replace $\mathbb{Z}$ by $\mathrm{End}(E_0)$ in the definition of $\Xi^{\mathbb{Z}\ell d}_{p,\delta q}$). Alternatively, one can appeal to Silverman's specialization theorem [1983]. □

To get the desired contradiction, it remains to show that the degrees

$$d_\xi = [k(\xi) : \mathbb{Q}]$$

too are bounded from above on the set Ξ.

2.1.2. *Heights of relations bounded by degrees.*

Lemma 4. *There exist two constants c, γ such that for any point $\xi \in \Xi$, there exist two integers $n \neq 0, m$ with $|n|, |m| \leq c d_\xi^\gamma$ such that $ns(\xi) - ms^R(\xi) = 0$.*

Proof. By [Bertrand et al. 2016], Corollary of Section 3.1, there exists a constant c' such that if $s(\xi)$ is a torsion point of G_ξ, its order n is bounded from above by $c' d_\xi^4$, so $(n, 0)$ satisfies the required condition. We can therefore assume that the Ribet point $s(\xi)$, hence $q(\xi)$ by Lemma 1, is not a torsion point. For $\xi \in \Xi$, there exist $a, b \in \mathbb{Z}$, not both 0, such that $ap(\xi) - b\delta q(\xi) = 0$, and since $q(\xi) \notin E_0^{\mathrm{tor}}$, any such relation will automatically imply $a \neq 0$. The points $p(\xi), \delta q(\xi)$ are defined over $k(\xi)$, and have heights $\leq c_0$. By works of Masser and David (see for instance Lemma 6.1 of [Barroero 2017]), there then exists such a relation with $max(|a|, |b|) \leq c_1 d_\xi^{\gamma_1}$ for some constants c_1, γ_1.

By our running hypothesis that $q(\xi)$ is not torsion, the set of such relations (trivial one included) is a free $\mathbb{Z}$ module of rank 1, and its generator (a_0, b_0) satisfies the above bound.

Consider now the nontorsion Ribet point $s(\xi)$ (so, $s^R(\xi)$ too is nontorsion), and let $(n_0 \neq 0, m_0) \in \mathbb{Z}^2$ be a generator of the group of relations $ns(\xi) - ms^R(\xi) = 0$, which is again free of rank 1. Projecting to E_0, we then have $n_0 p(\xi) - m_0 \delta q(\xi) = 0$. So, there exists $d \in \mathbb{N}$ such that $(n_0, m_0) = d.(a_0, b_0)$, and $a_0 s(\xi) - bs_0^R(\xi)$ is a torsion point of $G_{q(\xi)}$, of exact order d since (n_0, m_0) is minimal. Since it projects to 0 on E_0, it is actually a d-th root of unity ζ_d. Now, both $s(\xi)$ and $s^R(\xi)$ are defined over $k(\xi)$ (since s and s^R are global sections of $G \to X$), so ζ_d too lies in $k(\xi)$. Since ζ_d has order d, this implies that $d \leq c_2 d_\xi^{\gamma_2}$, say with $\gamma_2 = 2$.

In conclusion, for any $\xi \in \Xi$, there is a linear relation $ns(\xi) - ms^R(\xi) = 0$, with $(n, m) \in \mathbb{Z}^2, n \neq 0$ and $max(|n|, |m|) \leq c d_\xi^\gamma$ for some constants c and $\gamma = \gamma_1 + \gamma_2$. □

2.1.3. *Counting relations of bounded height.* In this step and the next one, we extend the scalars from $\mathbb{Q}^{\mathrm{alg}}$ to $\mathbb{C}$, but still write $X, K = \mathbb{C}(X)$, etc, instead of $X_{\mathbb{C}}, K \otimes \mathbb{C}, \ldots$. We sometimes indicate by the exponent $^{\mathrm{an}}$ the analytic object attached to an algebraic one over $\mathbb{C}$.

We now follow the usual procedure of studying the lifts to a universal covering of the relations considered in Lemma 4, and bounding their number via (generalizations of) the Pila–Wilkie theorem for a relevant o-minimal structure. There are several ways to implement this method. For instance, we can

(A) choose a fundamental domain $\mathcal{F}$ for the uniformization map unif : $\tilde{G} \simeq \mathbb{C} \rtimes (\mathbb{C} \times \tilde{X}) \to G^{\mathrm{an}}$, and count the relations in $\tilde{G}$ when the transcendence degree over $\mathbb{C}$ of the field of definition of $(\mathrm{unif}_{|\mathcal{F}})^{-1} \circ s$ is large enough. Here, $\mathcal{F}$ is unbounded, but by work of Peterzil and Starchenko, a convenient choice allows to work in the o-minimal structure $\mathbb{R}_{an,\exp}$; or

(B) fix a simply connected domain $D \subset X^{\mathrm{an}}$, consider the exponential morphism $\exp_G$, restricted over D, and count the relations in $(\mathrm{Lie}\, G)/D \simeq (\mathbb{C} \rtimes \mathbb{C}) \times D$ when the transcendence degree over $\mathbb{C}(X)$ of the field of definition of $\exp_G^{-1}(s_{|D})$ is sufficiently large. Here, D can be compact, and it suffices to work in the o-minimal structure $\mathbb{R}_{\mathrm{an}}$.

An advantage of (A) is its impact on effectivity, as alluded to in Comment (vii) of Section 1.3 (see also Remark 3 of Section 4.3). But as in [Bertrand et al. 2016, §3.3], we here follow the more elementary approach (B), taking advantage of the computation of transcendence degrees already established in this paper.

So, let (D, ξ_0) be a pointed set in X^{an}, homeomorphic to a closed disk. The group scheme G/X defines an analytic family G^{an} of Lie groups over the Riemann surface X^{an}. Similarly, its relative Lie algebra $(\mathrm{Lie}\, G)/X$ defines an analytic vector bundle $\mathrm{Lie}\, G^{\mathrm{an}}$ over X^{an}, of rank 2. We denote by Π_G the $\mathbb{Z}$-local system of periods of $G^{\mathrm{an}}/X^{\mathrm{an}}$; it is the kernel of the exponential exact sequence of analytic sheaves over X^{an}:

$$0 \to \Pi_G \to \mathrm{Lie}\, G^{\mathrm{an}} \xrightarrow{\exp_G} G^{\mathrm{an}} \to 0.$$

For any U_0 in $\mathrm{Lie}(G_{\xi_0}(\mathbb{C}))$ such that $\exp_{G_{\xi_0}}(U_0) = s(\xi_0) \in G_{\xi_0}(\mathbb{C})$, there exists a unique analytic section U of $\mathrm{Lie}(G^{\mathrm{an}})/D$ (meaning over a neighborhood of D), such that

$$U(\xi_0) = U_0 \quad \text{and} \quad \forall \xi \in D,\ \exp_{G_\xi^{\mathrm{an}}}(U(\xi)) = s(\xi).$$

Since D is fixed, we will just write $U = \log_G(s)$, although only its class modulo Π_G is well defined. Similarly, let $U^R = \log_G(s^R)$ for the Ribet section s^R. By the same process for $E_{0/X}$ (and the tacit assumption that the logarithms at ξ_0 are chosen in a compatible way), the projection $p = \pi \circ s \in E_0(X)$ admits as logarithm $\log_{E_0}(p) := u = d\pi(U)$; we also set $v = \log_{E_0}(q)$, so $d\pi(U^R) := u^R = \delta v$.

We will use the explicit expressions given in [Bertrand et al. 2016] for U, U^R and Π_G. These hold on any simply connected domain of X^{an} where u, v and $u+v$ do not assume period values. This is ensured by the hypothesis, made at the beginning of Section 2.1, that p, q and $p+q$ vanish nowhere on X.

Let $K = \mathbb{C}(X)$ be the field of rational functions of X. Since $\mathrm{Lie}\, G$ is a vector bundle over X, it makes sense to speak of the field of definition $K(U)$ of U over K. Similarly, let $F_G = K(\Pi_G)$ be the field of definition of Π_G. Notice that the field $F_G(U)$ now depends only on the section s. Moreover, for the Ribet section s^R, we have:

Lemma 5. *The field of definition $F^R = K(U^R)$ of any logarithm U^R of s^R coincides with the field of periods F_G of G.*

Proof. The explicit expressions of Π_G and U^R given in [Bertrand et al. 2016, §A.1], show that both fields coincide with the field $K(v, \zeta(v))$, where ζ denotes the Weierstrass zeta function of the elliptic curve E_0. □

For any real number $T \geq 1$, set $\mathbb{Z}[T] = \{n \in \mathbb{Z}, |n| \leq T\}$, and consider the subset

$$\Xi[T] := \{\xi \in X(\mathbb{Q}^{\mathrm{alg}}) \mid \exists (n, m) \in (\mathbb{Z}[T])^2,\ n \neq 0,\ ns(\xi) - ms^R(\xi) = 0\}$$

of $\Xi = \Xi_s$. We then have:

Proposition 1. *Let D be a closed disk in X^{an}. For any $\epsilon > 0$, there exists a real number C_ϵ, depending only on X, E_0, q, s, D and ϵ, such that*

(a) *either, for any $T \geq 1$, there are at most $C_\epsilon T^\epsilon$ points in $D \cap \Xi[T]$; or*

(b) *the field $F_G(U)$ has transcendence degree at most 1 over the field F_G.*

The proof of Proposition 1 is given in Section 2.2 below, as a corollary of Habegger and Pila's "semirational" count [2016, Corollary 7.2].

2.1.4. *Logarithmic Ax.* Assume that conclusion (b) of Proposition 1 holds. Since $u = d\pi(U)$, the field $F_G(U)$ has transcendence degree at most 1 over $F_G(u)$, and

(b1) either u is algebraic over $F_G = K(v, \zeta(v))$, in which case we know by the Ax–Schanuel theorem on the universal vectorial extension of the elliptic curve E_0 (see for instance [Bertrand et al. 2016, §6, Case (SC3)]) that p and q are linearly dependent over $\mathrm{End}(E_0)$ modulo constants; or

(b2) $U = \log_G(s)$ is algebraic over $F_G(u)$, hence over $K(u, \zeta(u), v, \zeta(v))$, in which case we know by [Bertrand et al. 2016, Lemma 5.1], that s is a translate of a Ribet section by a constant one, i.e., one in $\mathbb{G}_m(\mathbb{C})$ since G is not isosplit. Then, $p = \pi \circ s$ and q are linearly dependent over $\mathrm{End}(E_0)$.

In both cases, we get a contradiction to our hypothesis that p and q are linearly independent over $\mathrm{End}(E_0)$ modulo $E_0(\mathbb{Q}^{\mathrm{alg}})$. So, conclusion (a) must hold.

2.1.5. *Conclusion.* It follows from Lemma 3 and a compactness argument (see [Masser and Zannier 2015, Lemma 8.2 and the paragraph after (9.2)]) that there exists a finite set of closed disks D_i in X^{an} and a constant c' such that the following holds: for any $\xi \in \Xi$, a positive proportion $\frac{1}{c'} d_\xi$ of the conjugates of ξ over k lie in one of the D_i's, say D_1. Now, all these conjugates are still in Ξ, since $\sigma(s^R(\xi)) = s^R(\sigma\xi)$ is a Ribet point of $G_{q(\sigma\xi)}$ for $\sigma \in Gal(\mathbb{Q}^{\mathrm{alg}}/k)$. Actually, by Lemma 4, all the conjugates of ξ over k lie in $\Xi[T]$ with $T = c d_\xi^\gamma$. Choosing $\epsilon = \frac{1}{2}\gamma$, we deduce from conclusion (a) that $D_1 \cap \Xi$ has at most $c'' d_\xi^{1/2}$ (and at least $\frac{1}{c'} d_\xi$) elements. Therefore, d_ξ is bounded from above on Ξ, and this concludes the proof of Theorem 1.w.

2.2. ***The semirational count.*** The proof of Proposition 1 uses Betti coordinates and maps, defined as follows. We recall that $D \subset X^{\mathrm{an}}$ is homeomorphic to a closed complex disk.

The sections of the local system Π_G over D form a $\mathbb{Z}$-module $\Pi_G(D) \subset \mathrm{Lie}\, G^{\mathrm{an}}(D)$ of rank 3, with a basis $\{\varpi_0, \varpi_1, \varpi_2\}$ such that ϖ_0 generates $\Pi_{\mathbb{G}_m}(D)$, and ϖ_1, ϖ_2 project to a basis ω_1, ω_2 of $\Pi_{E_0}(D)$. Then, any logarithm $U := \log_G(s)$ of a section s of G/X over the disk D can uniquely be written as

$$U = b_0 \varpi_0 + b_1 \varpi_1 + b_2 \varpi_2,$$

where b_0, b_1, b_2 are real analytic functions on D, with values in $\mathbb{C}$ for b_0, and in $\mathbb{R}$ for b_1 and b_2. We call (b_0, b_1, b_2) the Betti coordinates of U, and define the Betti map attached to U as

$$U_B = (b_0; b_1, b_2) : D \to \mathbb{C} \times \mathbb{R}^2,$$

Similarly, we write $U_B^R = (b_0^R; b_1^R, b_2^R)$ for the Betti map attached to $U^R = \log_G(s^R)$, and denote by $\mathcal{S}$ the image of the disk D under the map

$$\mathcal{U}_B := (U_B, U_B^R) : D \twoheadrightarrow \mathcal{S} \subset \mathbb{R}^4 \times \mathbb{R}^4 = \mathbb{R}^8.$$

We will work in the o-minimal structure $\mathbb{R}_{\mathrm{an}}$ of globally subanalytic sets.

Lemma 6. *$\mathcal{S} = \mathcal{U}_B(D)$ is a compact 2-dimensional set, definable in the structure $\mathbb{R}_{\mathrm{an}}$.*

Proof. By definition (or by inspection of the formulae in [Bertrand et al. 2016]), the maps U_B and U_B^R extend to real analytic maps on a neighborhood of the compact disk D. Therefore, $\mathcal{S} = \mathcal{U}_B(D)$ is a compact definable set. Furthermore, the Betti map $\pi \circ U_B^R := u_B^R = (b_1^R, b_2^R)$ attached to $u^R = \log_{E_0}(p^R)$ is an immersion (since $p^R = \delta q \in E_0(X)$ is not a constant section), so $\mathcal{S}$ is indeed a real surface. □

With this notation in mind, a point ξ of D lies in $D \cap \Xi$ if and only if

$$\exists(\nu \neq 0, \mu) \in \mathbb{Z}^2 \mid \exists(\beta_0, \beta_1, \beta_2) \in \mathbb{Z}^3,\ \nu U(\xi) - \mu U^R(\xi) = \beta_0 \varpi_0(\xi) + \beta_1 \varpi_1(\xi) + \beta_2 \varpi_2(\xi),$$

or alternatively, in terms of the Betti maps,

$$\exists(\nu \neq 0, \mu) \in \mathbb{Z}^2 \mid \exists(\beta_0, \beta_1, \beta_2) \in \mathbb{Z}^3,\ \nu U_B(\xi) - \mu U_B^R(\xi) = (\beta_0; \beta_1, \beta_2) \in \mathbb{Z} \times \mathbb{Z}^2 \subset \mathbb{C} \times \mathbb{R}^2.$$

Remark that:

- *If $|\nu|, |\mu|$ are bounded by some number T, then $|\beta_0|, |\beta_1|, |\beta_2| \leq C_1 T$* for some constant C_1, since D is compact.
- *Given any* real *numbers $\nu \neq 0, \mu, \beta_0, \beta_1, \beta_2$*, there are only *finitely many ξ's in D such that $\nu U_B(\xi) - \mu U_B^R(\xi) = (\beta_0; \beta_1, \beta_2)$*. Otherwise, $\nu u - \mu \delta v$ would be constant on D, contradicting the Ax–Schanuel theorem invoked in Section 2.1.4(b1).

We can now describe the definable set $\mathcal{Z}$ to which Habegger and Pila's semirational count [2016] will be applied. On the one hand, we have the affine space $\mathbb{R}^5$ with real coordinates $(\nu, \mu, \beta_0, \beta_1, \beta_2)$; we will indicate by the index $_*$ the complement of the hyperplane $\nu = 0$. On the other hand, we have the affine space $\mathbb{C} \times \mathbb{R}^2 = \mathbb{R}^4$ and its square $\mathbb{R}^8$, which is the target space of the map $\mathcal{U}_B$. We consider the incidence variety $\mathcal{Z}$ in $\mathbb{R}^5 \times \mathbb{R}^8$, with projections π_1 to $\mathbb{R}^5_* \subset \mathbb{R}^5$ and π_2 to $\mathcal{S} = \mathcal{U}_B(D) \subset \mathbb{R}^8$:

$$\mathcal{Z} = \{((\nu, \mu, \beta_0, \beta_1, \beta_2); (w := (w_0; w_1, w_2), w^R := (w_0^R; w_1^R, w_2^R))) \in \mathbb{R}^5 \times \mathcal{S} \subset \mathbb{R}^5 \times \mathbb{R}^8,$$
$$\text{such that } \nu \neq 0 \text{ and } \nu.w - \mu.w^R = (\beta_0; \beta_1, \beta_2) \in \mathbb{R} \times \mathbb{R}^2 \subset \mathbb{C} \times \mathbb{R}^2 = \mathbb{R}^4\}$$

By Lemma 6, $\mathcal{Z}$ is a definable subset of $\mathbb{R}^{13}$. Furthermore, $\mathcal{U}_B(D \cap \Xi) = \pi_2(\pi_1^{-1}(\mathbb{Z}^5_*))$.

Let $\epsilon \in \mathbb{R}_{>0}$. Given $T \geq 1$, let $\mathcal{Z}[T]$ be the subset $\pi_1^{-1}((\mathbb{Z}[T])^5_*)$ formed by those elements of $\mathcal{Z}$ whose projection to $\mathbb{R}^5_*$ have integer coordinates of height $\leq T$. By [Habegger and Pila 2016, Corollary 7.2], (with no $\mathbb{R}^\ell$), there is a constant C'_ϵ such that one of the following holds:

(a') $\pi_2(\mathcal{Z}[T]) \subset \mathcal{U}_B(D \cap \Xi[T]) \subset \mathcal{S}$ has less than $C'_\epsilon T^\epsilon$ elements. Recalling the two remarks above, we then deduce from an o-minimal uniformity argument (or from a zero estimate as in [Bertrand et al. 2016, Proposition 3.3]) that for some constant C_ϵ, there are at most $C_\epsilon T^\epsilon$ points $\xi \in D \cap \Xi$ for which $\nu U(\xi) - \mu U^R(\xi) \in \Pi_{G_\xi}$ for some $(\nu \neq 0, \mu) \in (\mathbb{Z}[T])^2$. This is conclusion (a) of Proposition 3.

(b') There is a definable connected curve $\mathcal{C} \subset \mathcal{Z}$ such that $\pi_1(\mathcal{C}) \subset \mathbb{R}^5_*$ is semialgebraic and $\pi_2(\mathcal{C}) \subset \mathcal{S}$ has (real) dimension 1. Let $\mathcal{T} \subset D \subset X(\mathbb{C})$ be the inverse image of $\pi_2(\mathcal{C})$ under the map $\mathcal{U}_B$. We can view $\mathcal{C}$ as parametrized by the curve $\mathcal{T}$. The coordinates $\mu, \nu, \beta_0, \beta_1, \beta_2; w_0, w_1, w_2, w_0^R, w_1^R, w_2^R$ on $\mathbb{R}^5 \times \mathbb{R}^8$, restricted to $\mathcal{C}$, then become functions of the (real) variable $\gamma \in \mathcal{T}$. Since $\pi_1(\mathcal{C})$ is semialgebraic, the functions $\mu(\gamma), \nu(\gamma), \beta_0(\gamma), \beta_1(\gamma), \beta_2(\gamma)$ generate a field of transcendence degree 1 (or 0, if constant) over $\mathbb{C}$. In view of the incidence relations, *whose ν-component does not vanish* by definition, the restrictions to $\mathcal{T}$ of the functions $w_0 = b_0, w_1 = b_1, w_2 = b_2$ generate a field of transcendence degree ≤ 1 over the field generated by the restrictions to $\mathcal{T}$ of the functions $w_0^R = b_0^R, w_1^R = b_1^R, w_2 = b_2^R$. Recalling that $U = b_0\varpi_0 + b_1\varpi_1 + b_2\varpi_2$, and similarly with U^R, we deduce that $U_{|\mathcal{T}}$ generate a field of transcendence degree ≤ 1 over the field generated by $U^R_{|\mathcal{T}}$ and the $\varpi_{i|\mathcal{T}}$'s. By complex analyticity, the corresponding algebraic relation extends to D, so U generates a field of transcendence degree ≤ 1 over the field $F^R.F_G$ generated over $\mathbb{C}(X)$ by U^R and the ϖ_i's. In view of Lemma 5, this is Conclusion (b), and the proof of Proposition 1 is completed. □

3. The weakly special case over $\mathbb{Z}$

From now on, we assume that *the sections p and q are linearly dependent over* End(E_0) *modulo the subgroup $E_0(\mathbb{Q}^{\text{alg}})$* of constant sections of $E_0(X)$, and look for a proof of Theorem 1. Since its statement is invariant under multiplication of s by a positive integer, and since q is not constant, we can assume without loss of generality that the generic relation they satisfy takes the form

$$p = \rho q + p_0, \quad \text{with } \rho \in \text{End}(E_0),\ p_0 \in E_0(\mathbb{Q}^{\text{alg}}),\ p_0 \notin E_0^{\text{tor}}(\mathbb{Q}^{\text{alg}})$$

(if p_0 is torsion, the conclusion of Theorem 1 is trivially satisfied). In such a case, the initial Step 2.1.1 of the previous proof simply does not hold; contrary to the situation of Lemma 5, the set

$$\Xi^{\mathbb{Z}\ell d}_{p,\delta q} = \{\xi \in X(\mathbb{Q}^{\text{alg}}) \mid p(\xi) \text{ and } \delta q(\xi) \text{ are linearly dependent over } \mathbb{Z}\}$$

may well have unbounded height.

In this section, we show that if

$$\rho = r \in \mathbb{Z}, \quad r \neq 0,$$

upper bounds for the height on $\Xi^{\mathbb{Z}\ell d}_{p,\delta q}$, hence on its subset Ξ, can still be recovered, thanks to Silverman's theorem and basic orthogonality properties of Néron–Tate pairings. Theorem 1 then follows by reproducing most of the previous proof.

3.1. ***Bounded height.*** Let again h denote the height on $X(\mathbb{Q}^{\text{alg}})$ attached to a divisor of degree 1.

Proposition 2. *Let $p, q \in E_0(X)$, $p_0 \in E_0(\mathbb{Q}^{\text{alg}})$, q not constant, and assume that there exists a nonzero integer r such that $p = rq + p_0$. Then, there exists a constant c_0' such that $h(\xi) \le c_0'$ for any $\xi \in \Xi^{\mathbb{Z}\ell d}_{p,\delta q}$, hence for any $\xi \in \Xi$.*

Proof. This follows from an elementary computation, using the fact that for any $\rho \in \text{End}(E_0)$, the Néron–Tate height of $\rho q(\xi)$ is $\rho\bar{\rho}$ times that of $q(\xi)$. The following argument is based solely on orthogonality properties. Assume for a contradiction that there exists a sequence ξ_n, $n \in \mathbb{N}$, of points of $\Xi^{\mathbb{Z}\ell d}_{p,\delta q}$ whose heights $h(\xi_n)$ tend to infinity. Denote by $\langle \cdot, \cdot \rangle_{\text{geo}}$ the (geometric) Néron–Tate pairing on $E_0(K^{\text{alg}}) \times E_0(K^{\text{alg}})$, where $K = \mathbb{Q}^{\text{alg}}(X)$, and by $\langle \cdot, \cdot \rangle_{\text{ari}}$ the (arithmetic) Néron–Tate pairing on $E_0(\mathbb{Q}^{\text{alg}}) \times E_0(\mathbb{Q}^{\text{alg}})$.

Recall that for both pairings, the adjoint of $\rho \in \text{End}(E_0)$ is its complex conjugate. In particular, $\delta q(\xi) = -\bar{\delta} q(\xi)$ is orthogonal to $q(\xi)$, so $\langle p(\xi_n), q(\xi_n)\rangle_{\text{ari}} = 0$ for all n. By Silverman [1983] (or see [Lang 1983, p. 306]), we deduce that

$$\langle p, q\rangle_{\text{geo}} = \lim_{n\to\infty} \frac{\langle p(\xi_n), q(\xi_n)\rangle_{\text{ari}}}{h(\xi_n)} = 0.$$

Now, $p = rq + p_0$, and the constant part $E_0(\mathbb{Q}^{\text{alg}})$ is orthogonal to the full space $E_0(K^{\text{alg}})$ for the geometric pairing. So

$$\langle p, q\rangle_{\text{geo}} = \langle rq, q\rangle_{\text{geo}} + \langle p_0, q\rangle_{\text{geo}} = r\langle q, q\rangle_{\text{geo}} \quad \text{with } r \neq 0.$$

Therefore, the section q has vanishing Néron–Tate height, hence must be constant, contrary to our hypothesis. □

3.2. ***Algebraic (in)dependence.*** Assuming that $p = rq + p_0$ as above, we now follow the proof of Section 2.1. All its steps go through, except that conclusion (b) of Proposition 1 is now automatically satisfied. Indeed, we have $u = rv + u_0$, where $u_0 \in \text{Lie}\, E_0(\mathbb{C})$ is a conveniently chosen elliptic logarithm of p_0, so $K(u)$ lies in the field $K(v) \subset F_G$, and automatically, $U = \log_G(s)$ generates a field of transcendence degree at most 1 over F_G.

To overcome this difficulty, we will now deduce from the generic relation $p = rq + p_0$ that Conclusion (b) can here be replaced by the more precise statement that

($\text{b}^\sharp$) the field $F_G(U)$ is algebraic over the field $F_G(u) = F_G$

(which is actually conclusion (b2) of Section 2.1.4).

To check this, we use the same incidence variety $\mathcal{Z}$ as in Section 2.2, and follow Alternative (b') of the discussion. Notice that any relation $\nu U(\xi) - \mu U^R(\xi) = \beta_0\varpi_0(\xi) + \beta_1\varpi_1(\xi) + \beta_2\varpi_2(\xi)$, projected to Lie E_0, yields $\nu u(\xi) - \mu u^R(\xi) = \beta_1\omega_1 + \beta_2\omega_1$ hence since $u^R = \delta v$:

$$(\nu r - \mu\delta)v(\xi) = \beta_1\omega_1 + \beta_2\omega_2 - \nu u_0.$$

Restricting this relation to the real curve $\mathcal{T} \subset D$, and recalling that $\nu \neq 0$, $r \neq 0$ and $\delta \notin \mathbb{R}$, we deduce that if Alternative (b') holds, then the field generated over $\mathbb{C}$ by the restriction of the function v to $\mathcal{T}$ lies in the field generated over $\mathbb{C}$ by the restriction to $\mathcal{T}$ of the real functions μ, ν and the β_i's, $i = 1, 2$. Since the latter field has transcendence degree at most 1 over $\mathbb{C}$, while v is not constant, the two fields have the same algebraic closure, in which u lies. The full incidence relation then implies that U is algebraic over the field $F^R.F_G(u) = F_G$. This is conclusion ($\mathrm{b}^\sharp$).

So, $\log_G(s)$ is algebraic over F_G. As explained in case (b2) of Section 2.1.4, Lemma 5.1 of [Bertrand et al. 2016] then implies that p and q are linearly dependent over $\mathrm{End}(E_0)$ and Theorem 1 is established in this "$\rho = r \in \mathbb{Z}$, $r \neq 0$-weakly special" case. □

4. End of proof of Theorem 1

4.1. ***From weakly special to constant.*** In this subsection, we assume that the projection $p \in E_0(X)$ of $s \in G(X)$ and the section $q \in E_0(X)$ are linked by a generic relation of arbitrary shape:

$$p = \rho q + p_0, \quad \text{with } \rho \in \mathrm{End}(E_0),\ p_0 \in E_0(\mathbb{Q}^{\mathrm{alg}}).$$

We will deduce from the previous section that either p and q are linearly dependent over $\mathrm{End}(E_0)$ (as predicted by Theorem 1), or we may assume that $\rho = 0$, i.e., p itself is a constant section.

Replacing s by $2s$ if necessary, we can write $\rho = r + r'\delta \in \mathbb{Z} \oplus \mathbb{Z}\delta \subset \mathrm{End}(E_0)$, and consider the basic Ribet section $s_{r'\alpha} = r's^R$ of $G = G_q$ over X. Its projection to $E_0(X)$ is the section $r'p^R = r'\delta q$. Therefore, the section $s' := s - s_{r'\alpha}$ of G/X projects to

$$\pi(s') := p' = p - r'\delta q = rq + p_0.$$

Moreover, for any $\xi \in X(\mathbb{Q}^{\mathrm{alg}})$, $s_{r'\alpha}(\xi) = s_{r'\alpha, q(\xi)}$ is by definition a Ribet point of $G_{q(\xi)}$. Consequently, the set $\Xi := \Xi_s$ of points of $X(\mathbb{Q}^{\mathrm{alg}})$ where $s(\xi)$ is a Ribet point coincides with the set $\Xi_{s'}$ similarly attached to s', which is therefore infinite. Since $r \in \mathbb{Z}$, we deduce from the result of Section 3 that either p' and q, hence p and q, are linearly dependent over $\mathrm{End}(E_0)$, or that $r = 0$.

Assume now that $r = 0$, so the generic relation reads: $p = r'\delta q + p_0$, and consider again the section $s' = s - r's^R$, which projects to $p' = p_0$. The corresponding set $\Xi_{s'}$ is still infinite. Therefore, we have reduced the proof of Theorem 1 to the case where $\rho = 0$, i.e., where the projection p of s is a constant section p_0. We must then show that p_0 is necessarily a torsion point.

4.2. ***The constant case.*** The word constant here refers not to the semiabelian scheme G/X, which we still assume to be nonconstant ($q \notin E_0(\mathbb{Q}^{\mathrm{alg}})$), but to the section $\pi \circ s := p = p_0 \in E_0(\mathbb{Q}^{\mathrm{alg}})$. However, the duality properties of the Poincaré biextension $\mathcal{P}_0$ of $E_0 \times \hat{E}_0$ by $\mathbb{G}_m$ enable us to permute the roles of q and p, thereby translating the problem into one on the constant semiabelian variety $G'_{p_0} = \mathcal{P}_{0|p_0 \times \hat{E}_0} \in \mathrm{Ext}(\hat{E}_0, \mathbb{G}_m)$ parametrized by the point p_0 of (the bidual of) E_0. We must then prove that p_0 is torsion, i.e., that G'_{p_0} is isosplit.

Assume for a contradiction that p_0 is not torsion. Then for each ξ in the set Ξ, there is a relation $np_0 - m\delta q(\xi) = 0$ with $nm \neq 0$, so $q(\xi)$ lies in the divisible hull of $\mathbb{Z}.\delta p_0$, and is not torsion either. Consider the constant semiabelian surface $G'_{p_0} \in \mathrm{Ext}(\hat{E}_0, \mathbb{G}_m)$. By duality, we can view s as a section $\check{s} \in G'_{p_0}(X)$, and $s(\xi)$ as a point $\check{s}(\xi)$ on G'_{p_0} projecting to $q(\xi)$ in $\hat{E}_0$. Furthermore, $\check{s}(\xi)$ is a nontorsion Ribet point of G'_{p_0} if and only if $s(\xi)$ is a nontorsion Ribet point of $G_{q(\xi)}$: in the setting of Section 1.2, this is clear when $\varphi - \hat{\varphi}$ is an isomorphism, and it remains true in general via an isogeny. (In fact, it is proven in [Bertrand and Edixhoven 2019, Remark 5.4.1], that the 1-motive attached to $s_{\varphi,q}$ is isogenous to its Cartier dual as soon as $\varphi - \hat{\varphi}$ is an isogeny.)

Therefore, the image $\check{s}(X)$ of $\check{s}$ is an irreducible curve in G'_{p_0} which contains infinitely many points of the group Γ'_0 formed by all the Ribet points of G'_{p_0}. Since this group has finite rank (at most 1), McQuillan's Mordell–Lang theorem [1995], as recalled in Section 1.3(v), can be applied to G'_{p_0}. We derive that $\check{s}$ factors through a translate by a Ribet point of a strict connected algebraic subgroup of G'_{p_0}. Since p_0 is not torsion, the only such one is $\mathbb{G}_m$, so $q(X)$ reduces to a point of $\hat{E}_0$. This contradicts our assumption that q is not constant, and concludes the proof of Theorem 1. □

4.3. ***Further comments.*** We here list properties of Ribet points and sections which although not used in the proof, may be relevant to further studies of unlikely intersections.

Remark 1 (in relation with Proposition 2). Attached to the divisor at infinity D_ξ of the standard compactification of $G_{q(\xi)}$, there is a canonical "relative height" $\hat{h}_{D_\xi}$, which vanishes on the Ribet points of $G_{q(\xi)}$; see [Bertrand 1995, §3]. Is there a Zimmer-like comparison of $\hat{h}_{D_\xi}$ with a Weil height h_{D_ξ}, of the type $\hat{h}_{D_\xi} - h_{D_\xi} = O((\hat{h}(q(\xi)))^{1/2})$, or even just $o(\hat{h}(q(\xi)))$, where $\hat{h}$ is the Néron–Tate height on $\hat{E}_0(\mathbb{Q}^{\mathrm{alg}})$? Bounded height on Ξ would then follow in all cases, "weakly special" or not. See [Chambert-Loir 1999, Theorem 5.5] for an Arakelov approach to this problem.

Remark 2 (on the Betti maps). Let $\xi \in \Xi$. By [Bertrand 1995, Theorem 4], the Ribet point $s(\xi)$ lies in the maximal compact subgroup of its fiber G_ξ^{an}. So its logarithm $U(\xi)$ lies in $\Pi_{G_\xi} \otimes \mathbb{R}$, and its Betti coordinate $b_0(\xi)$ is a real number. Similarly, the Betti coordinate b_0^R of the Betti map U_B^R attached to $U^R = \log_G(s^R)$ is actually real-valued. But a priori, not the Betti coordinate b_0 of U. It would be interesting to characterize the sections $s \in G(X)$ whose images meet the union of the maximal compact subgroups of all the fibers infinitely often.

Remark 3 (about effectivity). As suggested in Section 2.1.3(A) (see also Section 1.3(vii)), making the "constants" of the text effective in terms of the initial datas X, E_0, q, s, requires a global version of Proposition 1. One should here start with the uniformization map $\mathrm{Unif}: \tilde{\mathcal{P}}_0 \simeq \mathbb{C} \rtimes (\mathbb{C} \times \mathbb{C}) \to \mathcal{P}_0^{\mathrm{an}}$ of the Poincaré biextension itself, thereby reflecting the symmetric roles played by p and q in the construction of Ribet sections. As far as the dependence in s is concerned, a first aim would be to show that these constants are uniformly bounded in terms of the degree of the curve $W = s(X)$ in a projective embedding of G. We point out that this aim has indeed been reached in various versions of the Mordell–Lang problem itself; see [Hrushovski and Pillay 2000] for a differential algebraic approach (inspired by work of Buium, and recently sharpened in [Binyamini 2017]) and [Rémond 2011, Thorem 2.4], for the general case.

5. The Zilber–Pink conjecture for $\mathcal{P}_0$

Pink's generalization of the conjectures on unlikely intersections proposed by Bombieri, Masser, Zannier and by Zilber asserts:

Conjecture [Pink 2005, Conjecture 1.3]. *Let $\mathcal{S}/\mathbb{C}$ be a mixed Shimura variety, and let W be an irreducible algebraic subvariety of $\mathcal{S}$, of dimension d. Assume that the intersection of W with the union of all the special subvarieties of $\mathcal{S}$ of codimension $> d$ is Zariski dense in W. Then, W is contained in a special subvariety of $\mathcal{S}$ of positive codimension.*

As in the text, let again $E_0/\mathbb{Q}^{\text{alg}}$ be an elliptic curve with complex multiplications, with dual $\hat{E}_0 \simeq \text{Ext}(E_0, \mathbb{G}_m)$, and let $\mathcal{P}_0/\mathbb{Q}^{\text{alg}}$ be the Poincaré biextension of $E_0 \times \hat{E}_0$ by $\mathbb{G}_m$. This is a $\mathbb{G}_m$-torsor over $E_0 \times \hat{E}_0$, which admits two families of group laws. Namely, for any $q \in \hat{E}_0$, the restriction of $\mathcal{P}_0$ above $E_0 \times \{q\}$ is the semiabelian variety attached to q, viewed as a point in $\text{Ext}(E_0, \mathbb{G}_m)$, while for any $p \in E_0$, the restriction of $\mathcal{P}_0$ above $\{p\} \times \hat{E}_0$ is the semiabelian variety attached to p, viewed by biduality as a point in $\text{Ext}(\hat{E}_0, \mathbb{G}_m) \simeq E_0$. The important point in this section is that $\mathcal{P}_0$ admits a canonical structure of a mixed Shimura variety, which is described in detail in [Bertrand and Edixhoven 2019]. However, only a minimal knowledge of MSV theory will be needed to prove Theorem 3 of the introduction.

Before proving this theorem, we note (as pointed out by J. Pila) that it completely establishes Theorem 4, i.e., Pink's conjecture for the MSV $\mathcal{S} = \mathcal{P}_0$ when the variety W is defined over $\mathbb{Q}^{\text{alg}}$. Indeed, if the dimension d of W is 0 or 3, there is nothing to prove. If $d = 2$, then the special subvarieties of $\mathcal{P}_0$ of codimension $> d$ are its special points, and the statement reduces to the André–Oort conjecture, which follows in this case from [Gao 2017b, Theorem 13.6]. So, only the case $d = 1$, i.e., Theorem 3, needs to be treated.

Through the first family of group laws above, the projection $\varpi : \mathcal{P}_0 \to \hat{E}_0$ turns $\mathcal{P}_0$ into the universal extension $\mathcal{G}$ of E_0 by $\mathbb{G}_m$, over the moduli space $\hat{E}_0$. For any integer n, we will denote by $[n]_{\mathcal{G}}$ the morphism of multiplication by n of the group scheme $\mathcal{G}/\hat{E}_0$. Its Ribet sections are well defined, and we call their images *Ribet curves of $\mathcal{P}_0$, in the sense of $\mathcal{G}/\hat{E}_0$*. Similarly, the projection $\varpi' : \mathcal{P}_0 \to E_0$ turns $\mathcal{P}_0$ into a group scheme $\mathcal{G}'/E_0$, with morphisms $[n]_{\mathcal{G}'}$ and *Ribet curves of $\mathcal{P}_0$, in the sense of $\mathcal{G}'/E_0$*. Furthermore, $[n]_{\mathcal{G}}$ and $[n]_{\mathcal{G}'}$ induce the same morphism $[n]$ on the fiber $\mathbb{G}_m$ of (ϖ, ϖ') above $(0, 0)$. With these definitions in mind, we have the following explicit necessary conditions for an irreducible curve to be special in $\mathcal{P}_0$. It follows from [Bertrand and Edixhoven 2019, §5], (see also [Bertrand 2011, §2]) that they are also sufficient, but we will not need this sharper result.

Proposition 3. *Let C be a special curve of the MSV $\mathcal{P}_0$. Then*:

(i) *If $\varpi : C \to \hat{E}_0$ is dominant, C is a Ribet curve in the sense of $\mathcal{G}/\hat{E}_0$.*

(ii) *If $\varpi' : C \to E_0$ is dominant, C is a Ribet curve in the sense of $\mathcal{G}'/E_0$.*

(iii) *If $(\varpi', \varpi)(C)$ is a point (p_0, q_0) of $E_0 \times \hat{E}_0$, this point is a torsion point, and C is the fiber of $\mathcal{P}_0$ above (p_0, q_0).*

Notice that most special curves C satisfy both (i) and (ii), and are therefore Ribet curves in both senses. This reflects the self-duality of nontorsion Ribet sections, already encountered in Section 4.2. As for (iii), it occurs if neither (i) nor (ii) are satisfied.

Proof. We will use the following facts, for which we refer to [Pink 2005; Gao 2017a]:

(F1) A point P of $\mathcal{P}_0$ is special (if and) only if $(p, q) = (\varpi', \varpi)(P)$ is torsion in $E_0 \times \hat{E}_0$ and P is torsion in the (isosplit) extension $\mathcal{G}_q$ (equivalently, in the isosplit $\mathcal{G}'_p$).

(F2) A special curve of $\mathcal{P}_0$ contains a Zariski-dense set of special points, hence by F1 a Zariski-dense set of torsion points of the various fibers of $\mathcal{G}/\hat{E}_0$ (or of $\mathcal{G}'/E_0$).

(F3) The image of a special subvariety under a Shimura morphism (such as ϖ', ϖ, $[n]_{\mathcal{G}}$, $[n]_{\mathcal{G}'}$) is a special subvariety.

Let then $C \subset \mathcal{P}_0 = \mathcal{G}$ be a special curve, dominating $\hat{E}_0$ as in (i). By base extension along the finite cover $\varpi : X := C \to \hat{E}_0$, we can view the diagonal map $X \to C_X$ as a section s of the group scheme $G = \mathcal{G}_X := \mathcal{G} \times_{\hat{E}_0} X$ over X. We can now apply Lemma 2 of Section 1.3 (relative Manin–Mumford) to $s \in G(X)$: by Facts F1 and F2, the set Ξ_s^{tor} is infinite and we infer that s is a Ribet section of G/X, or factors through a torsion translate of $\mathbb{G}_{m/X} = \mathbb{G}_m \times X$. In the first case, the image $C \subset \mathcal{G}$ of $s(X) \subset C_X \subset \mathcal{G}_X$ is a Ribet curve of $\mathcal{P}_0$ in the sense of $\mathcal{G}/\hat{E}_0$, as was to be shown.

In the second case, a multiple $C' := [n]_{\mathcal{G}}(C)$ of C lies in the fiber $\mathbb{G}_m \times \hat{E}_0$ of $\mathcal{P}_0$ above $p = 0$, and is still a special curve of $\mathcal{P}_0$ by F3. So, by F2, C' contains infinitely many special points of $\mathcal{P}_0$ lying in $\mathbb{G}_m \times \hat{E}_0$. But by F1, these special points are contained in (in fact, fill up) the torsion of the group $\mathbb{G}_m \times \hat{E}_0$. We can now apply the standard Manin–Mumford theorem [Hindry 1988] to $C' \cap (\mathbb{G}_m \times \hat{E}_0)^{\text{tor}}$, and deduce that C' is a torsion translate of $\mathbb{G}_m \times \{0\}$ or of $\{1\} \times \hat{E}_0$. The first conclusion cannot occur since C' too dominates $\hat{E}_0$. So, a multiple $[m]C' = [mn]_{\mathcal{G}}(C)$ of C is the image of the unit section of $\mathcal{G}/\hat{E}_0$. Therefore, C is in all cases a Ribet curve of $\mathcal{P}_0$ in the sense of $\mathcal{G}/\hat{E}_0$.

The same proof applies to (ii), while (iii) easily follows from F1 (or from F3, in view of [Gao 2017a]). This concludes the proof of Proposition 3. □

We can now turn to the proof of Theorem 3. We will need the following complement to Fact F3:

(F4) Under a Shimura morphism, the irreducible components of the inverse image of a special subvariety are special subvarieties.

Proof of Theorem 3. Let $W/\mathbb{Q}^{\text{alg}}$ be an irreducible algebraic curve in $\mathcal{P}_0$, which contains infinitely many points lying on special curves of $\mathcal{P}_0$. We must show that W is contained in a special surface of $\mathcal{P}_0$. We deduce from Proposition 3 that

(a) W contains infinitely many points lying on Ribet curves in the sense of $\mathcal{G}/\hat{E}_0$, and if not,

(b) W contains infinitely many points lying on Ribet curves in the sense of $\mathcal{G}'/E_0$, and if not,

(c) W contains infinitely many points lying in the fibers of $\mathcal{P}_0$ above the torsion points of $E_0 \times \hat{E}_0$.

Assume first that $\varpi : W \to \hat{E}_0$ is dominant, and that we are in case (a). Base changing along $\varpi : X = W \to \hat{E}_0$ as above, we may view the diagonal map $X \to W_X \subset \mathcal{G}_X = G$ as a section $s \in G(X)$, to which Theorem 1 (or the relative Mordell–Lang Theorem 2) of Section 1.1 applies. By (a), the set Ξ_s is infinite, and we infer that the sections p and q attached to s are linearly dependent over $\mathrm{End}(E_0)$. So, $(\varpi', \varpi)(W)$ lies in a torsion translate of an elliptic curve $B \subset E_0 \times \hat{E}_0$ passing through 0. By [Gao 2017a], these are special curves of the MSV $E_0 \times \hat{E}_0$. Therefore, by F4, W lies in a special surface of $\mathcal{P}_0$. Vice versa, the same conclusion holds if $\varpi' : W \to E_0$ is dominant and we are in case (b).

Secondly, assume that W still dominates $\hat{E}_0$, but that we are in case (b). As just pointed out, we can then assume that W does not dominate E_0, and so, projects to a point $p \in E_0$ under ϖ'. If p is not torsion, W lies in the nonisosplit extension $\mathcal{G}'_p = \varpi'^{-1}(p)$ (which is then not a special surface of $\mathcal{P}_0$). Now, the Ribet curves in the sense of $\mathcal{G}'/E_0$ meet $\mathcal{G}'_p$ at Ribet points of $\mathcal{G}'_p$, so by (b), W contains infinitely many Ribet points of $\mathcal{G}'_p$. We deduce from the standard Mordell–Lang theorem [McQuillan 1995] that W lies in a translate of $\mathbb{G}_m$ by a Ribet point. But then, W cannot dominate $\hat{E}_0$. So, p is a torsion point, and W lies in $\varpi'^{-1}(p)$, which is a special surface of $\mathcal{P}_0$ by F4. Vice versa, the same conclusion holds if $\varpi' : W \to E_0$ is dominant and we are in case (a).

Thirdly, assume that W dominates $\hat{E}_0$ or E_0, and that we are in case (c). Then, the projection W' of W in $E_0 \times \hat{E}_0$ is a curve which contains infinitely many torsion points of $E_0 \times \hat{E}_0$. By Manin–Mumford, we deduce that W' lies in a torsion translate of an elliptic curve $B \subset E_0 \times \hat{E}_0$ passing through 0. So, W lies in a special surface of $\mathcal{P}_0$.

It remains to study the case when W projects to a point (p, q) of $E_0 \times \hat{E}_0$ under (ϖ', ϖ). Then, the only special curve of type (c) which meets W is the closure of W itself, so in case (c), (p, q) is a torsion point, and W lies in (many) a special surface of $\mathcal{P}_0$. Assume finally that we are in case (a), or in case (b). Then, W contains a Ribet point of $\mathcal{G}_q$, or of $\mathcal{G}'_p$, projecting to $p \in E_0$, or to $q \in \hat{E}_0$. In both cases, we deduce that the points p and q are linearly dependent over $\mathrm{End}(E_0)$. So, the projection to $E_0 \times \hat{E}_0$ of W lies in a torsion translate of an elliptic curve B passing through 0, and W lies in a special surface of $\mathcal{P}_0$. This concludes the proof of Theorem 3. □

Acknowledgements

Both authors thank the Fields Institute for invitations to their 2017 program on Unlikely Intersections, Heights, and Efficient Congruencing, where this work was initiated. We also thank Gareth Jones and Jonathan Pila for motivating discussions. Schmidt would further like to thank the Engineering and Physical Sciences Research Council for support under grant EP/N007956/1.

References

[Barroero 2017] F. Barroero, "CM relations in fibered powers of elliptic families", *J. Inst. Math. Jussieu* (online publication August 2017).

[Barroero and Capuano 2018] F. Barroero and L. Capuano, "Unlikely intersections in families of abelian varieties and the polynomial Pell equation", preprint, 2018. arXiv

[Bertrand 1995] D. Bertrand, "Minimal heights and polarizations on group varieties", *Duke Math. J.* **80**:1 (1995), 223–250. MR Zbl

[Bertrand 2011] D. Bertrand, "Special points and Poincaré bi-extensions", preprint, 2011. arXiv

[Bertrand 2013] D. Bertrand, "Unlikely intersections in Poincaré biextensions over elliptic schemes", *Notre Dame J. Form. Log.* **54**:3-4 (2013), 365–375. MR Zbl

[Bertrand and Edixhoven 2019] D. Bertrand and B. Edixhoven, "Pink's conjecture on unlikely intersections and families of semi-abelian varieties", preprint, 2019. arXiv

[Bertrand et al. 2016] D. Bertrand, D. Masser, A. Pillay, and U. Zannier, "Relative Manin–Mumford for semi-Abelian surfaces", *Proc. Edinb. Math. Soc.* (2) **59**:4 (2016), 837–875. MR Zbl

[Binyamini 2017] G. Binyamini, "Bezout-type theorems for differential fields", *Compos. Math.* **153**:4 (2017), 867–888. MR Zbl

[Chambert-Loir 1999] A. Chambert-Loir, "Géométrie d'Arakelov et hauteurs canoniques sur des variétés semi-abéliennes", *Math. Ann.* **314**:2 (1999), 381–401. MR Zbl

[Gao 2017a] Z. Gao, "A special point problem of André–Pink–Zannier in the universal family of Abelian varieties", *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5) **17**:1 (2017), 231–266. MR Zbl

[Gao 2017b] Z. Gao, "Towards the André–Oort conjecture for mixed Shimura varieties: the Ax–Lindemann theorem and lower bounds for Galois orbits of special points", *J. Reine Angew. Math.* **732** (2017), 85–146. MR Zbl

[Habegger and Pila 2016] P. Habegger and J. Pila, "O-minimality and certain atypical intersections", *Ann. Sci. Éc. Norm. Supér.* (4) **49**:4 (2016), 813–858. MR Zbl

[Hindry 1988] M. Hindry, "Autour d'une conjecture de Serge Lang", *Invent. Math.* **94**:3 (1988), 575–603. MR Zbl

[Hrushovski and Pillay 2000] E. Hrushovski and A. Pillay, "Effective bounds for the number of transcendental points on subvarieties of semi-abelian varieties", *Amer. J. Math.* **122**:3 (2000), 439–450. MR Zbl

[Jacquinot and Ribet 1987] O. Jacquinot and K. A. Ribet, "Deficient points on extensions of abelian varieties by $\mathbb{G}_m$", *J. Number Theory* **25**:2 (1987), 133–151. MR Zbl

[Jones and Schmidt 2017] G. Jones and H. Schmidt, "Pfaffian definitions of Weierstrass elliptic functions", preprint, 2017. arXiv

[Jones and Schmidt $\geq$ 2019] G. Jones and H. Schmidt, "Effective relative Manin–Mumford for families of $\mathbb{G}_m$-extensions of an elliptic curve", in preparation.

[Jones and Thomas 2018] G. O. Jones and M. E. M. Thomas, "Effective Pila–Wilkie bounds for unrestricted Pfaffian surfaces", preprint, 2018. arXiv

[Lang 1983] S. Lang, *Fundamentals of Diophantine geometry*, Springer, 1983. MR Zbl

[Masser and Zannier 2015] D. Masser and U. Zannier, "Torsion points on families of simple abelian surfaces and Pell's equation over polynomial rings", *J. Eur. Math. Soc. (JEMS)* **17**:9 (2015), 2379–2416. MR Zbl

[McQuillan 1995] M. McQuillan, "Division points on semi-abelian varieties", *Invent. Math.* **120**:1 (1995), 143–159. MR Zbl

[Pink 2005] R. Pink, "A common generalization of the conjectures of André–Oort, Manin–Mumford, and Mordell–Lang", preprint, 2005, Available at https://tinyurl.com/rpinkaom.

[Rémond 2011] G. Rémond, "Une remarque de dynamique sur les variétés semi-abéliennes", *Pacific J. Math.* **254**:2 (2011), 397–406. MR Zbl

[Silverman 1983] J. H. Silverman, "Heights and the specialization map for families of abelian varieties", *J. Reine Angew. Math.* **342** (1983), 197–211. MR Zbl

[Viada 2003] E. Viada, "The intersection of a curve with algebraic subgroups in a product of elliptic curves", *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5) **2**:1 (2003), 47–75. MR Zbl

Communicated by Jonathan Pila
Received 2018-11-06 Revised 2019-04-08 Accepted 2019-05-14

daniel.bertrand@imj-prg.fr *Institut de Mathematiques de Jussieu, Sorbonne Université, Paris, France*

harry.schmidt@unibas.ch *Mathematik und Informatik, Universität Basel, Switzerland*

mathematical sciences publishers msp

ALGEBRA AND NUMBER THEORY 13:6 (2019)
dx.doi.org/10.2140/ant.2019.13.1475

Congruences of parahoric group schemes

Radhika Ganapathy

Let F be a nonarchimedean local field and let T be a torus over F. With $\mathcal{T}^{NR}$ denoting the Néron–Raynaud model of T, a result of Chai and Yu asserts that the model $\mathcal{T}^{NR} \times_{\mathfrak{O}_F} \mathfrak{O}_F/\mathfrak{p}_F^m$ is canonically determined by $(\mathrm{Tr}_l(F), \Lambda)$ for $l \gg m$, where $\mathrm{Tr}_l(F) = (\mathfrak{O}_F/\mathfrak{p}_F^l, \mathfrak{p}_F/\mathfrak{p}_F^{l+1}, \epsilon)$ with ϵ denoting the natural projection of $\mathfrak{p}_F/\mathfrak{p}_F^{l+1}$ on $\mathfrak{p}_F/\mathfrak{p}_F^l$, and $\Lambda := X_*(T)$. In this article we prove an analogous result for parahoric group schemes attached to facets in the Bruhat–Tits building of a connected reductive group over F.

1. Introduction

Let F be a nonarchimedean local field, $\mathfrak{O}_F$ its ring of integers, and $\mathfrak{p}_F$ its maximal ideal. Let T be a torus over F. Such a torus is canonically determined by the lattice $\Lambda := X_*(T)$ together with the action of $\Gamma_F = \mathrm{Gal}(F_s/F)$ on it (here F_s is a separable closure of F). For large m, the action of Γ_F on Λ factors through the quotient Γ_F/I_F^m of Γ_F, where I_F^m is the m-th higher ramification subgroup (with upper numbering) of the inertia group I_F. This Galois group depends only on truncated data $\mathrm{Tr}_m(F) := (\mathfrak{O}_F/\mathfrak{p}_F^m, \mathfrak{p}_F/\mathfrak{p}_F^{m+1}, \epsilon)$, where ϵ is the natural projection of $\mathfrak{p}_F/\mathfrak{p}_F^{m+1}$ on $\mathfrak{p}_F/\mathfrak{p}_F^m$, via Deligne's theory; see (b) below.

Let $\mathcal{T}^{NR}$ denote the Néron–Raynaud model of T (see [Bosch et al. 1990]). The main result of [Chai and Yu 2001] asserts that $\mathcal{T}^{NR} \times_{\mathfrak{O}_F} \mathfrak{O}_F/\mathfrak{p}_F^m$ is canonically determined by $(\mathrm{Tr}_l(F), \Lambda)$ for $l \gg m$ (see Theorem 8.5 of [Chai and Yu 2001] for the precise statement; the parameters that l depends on are also explicitly determined there). With $\mathcal{T}$ denoting the neutral component of $\mathcal{T}^{NR}$ this also implies that $\mathcal{T} \times_{\mathfrak{O}_F} \mathfrak{O}_F/\mathfrak{p}_F^m$ is canonically determined by $(\mathrm{Tr}_l(F), \Lambda)$ with l as above. From the point of view of Bruhat–Tits theory, when the connected reductive group is a torus, the model $\mathcal{T}$ can be thought of as its Iwahori (or parahoric) group scheme. The purpose of this article is to prove an analogous result for parahoric group schemes attached to facets in the Bruhat–Tits building of a connected reductive group over F.

Our motivation for proving such a result arises naturally from the question of generalizing Kazhdan's theory of studying representation theory of split p-adic groups over close local fields to general connected reductive groups. Let us briefly recall the Deligne–Kazhdan correspondence:

(a) Given a local field F' of characteristic p and an integer $m \geq 1$, there exists a local field F of characteristic 0 such that F' is m-close to F, i.e., $\mathfrak{O}_F/\mathfrak{p}_F^m \cong \mathfrak{O}_{F'}/\mathfrak{p}_{F'}^m$.

MSC2010: primary 22E50; secondary 11F70.
Keywords: parahoric, close local fields.

(b) Deligne [1984] proved that if $\mathrm{Tr}_m(F) \cong \mathrm{Tr}_m(F')$, then the Galois groups $\mathrm{Gal}(F_s/F)/I_F^m$ and $\mathrm{Gal}(F'_s/F')/I_{F'}^m$ are isomorphic. This gives a bijection

$$\{\text{Iso. classes of cont., complex, f.d. representations of } \mathrm{Gal}(F_s/F) \text{ trivial on } I_F^m\}$$
$$\longleftrightarrow \{\text{Iso. classes of cont., complex, f.d. representations of } \mathrm{Gal}(F'_s/F') \text{ trivial on } I_{F'}^m\}.$$

Moreover, all of the above holds when $\mathrm{Gal}(F_s/F)$ is replaced by W_F, the Weil group of F.

(c) Let G be a split, connected reductive group defined over $\mathbb{Z}$. For an object X associated to the field F, we will use the notation X' to denote the corresponding object over F'. Kazhdan [1986] proved that given $m \geq 1$, there exists $l \geq m$ such that if F and F' are l-close, then there is an algebra isomorphism $\mathrm{Kaz}_m : \mathcal{H}(G(F), K_m) \to \mathcal{H}(G(F'), K'_m)$, where K_m is the m-th usual congruence subgroup of $G(\mathfrak{O}_F)$. Hence, when the fields F and F' are sufficiently close, we have a bijection

$$\{\text{Iso. classes of irr. admissible representations } (\Pi, V) \text{ of } G(F) \text{ such that } \Pi^{K_m} \neq 0\}$$
$$\longleftrightarrow \{\text{Iso. classes of irr. admissible representations } (\Pi', V') \text{ of } G(F') \text{ such that } \Pi'^{K'_m} \neq 0\}.$$

These results suggest that, if one understands the representation theory of $\mathrm{Gal}(F_s/F)$ for all local fields F of characteristic 0, then one can use it to understand the representation theory of $\mathrm{Gal}(F'_s/F')$ for a local field F' of characteristic p, and similarly, with an understanding of the representation theory of $G(F)$ for all local fields F of characteristic 0, one can study the representation theory of $G(F')$, for F' of characteristic p. This method has proved useful for studying the local Langlands correspondence for reductive p-adic groups in characteristic p via the corresponding theory in characteristic 0 (see [Badulescu 2002; Lemaire 2001; Ganapathy 2015; Aubert et al. 2016; Ganapathy and Varma 2017]). An obvious observation, that goes into proving the Kazhdan isomorphism, is

$$G(\mathfrak{O}_F)/K_m \cong G(\mathfrak{O}_F/\mathfrak{p}_F^m) \cong G(\mathfrak{O}_{F'}/\mathfrak{p}_{F'}^m) \cong G(\mathfrak{O}_{F'})/K'_m \tag{1-1}$$

if the fields F and F' are m-close.

A useful variant of the Kazhdan isomorphism is now available for split reductive groups. Let I be the standard Iwahori subgroup of G. It is shown in [Bruhat and Tits 1984] that there is a smooth affine group scheme $\mathcal{I}$ defined over $\mathfrak{O}_F$ with generic fiber $G \times_{\mathbb{Z}} F$ such that $\mathcal{I}(\mathfrak{O}_F) = I$. Define $I_m := \mathrm{Ker}(\mathcal{I}(\mathfrak{O}_F) \to \mathcal{I}(\mathfrak{O}_F/\mathfrak{p}_F^m))$. In Section 3 of [Ganapathy 2015], a presentation has been written down for this Hecke algebra $\mathcal{H}(G, I_m)$ (extending Theorem 2.1 of [Howe 1985] for GL_n). Furthermore if the fields F and F' are m-close, an argument of J.K. Yu (see Section 3.4.A of [Ganapathy 2015]) gives an isomorphism

$$\beta : I/I_m \to I'/I'_m. \tag{1-2}$$

Let us note here that unlike (1-1), the above isomorphism is not obvious since the group scheme $\mathcal{I}$ is defined over $\mathfrak{O}_F$ and not over $\mathbb{Z}$. In fact the above isomorphism is obtained by proving that the reduction $\mathcal{I} \times_{\mathfrak{O}_F} \mathfrak{O}_F/\mathfrak{p}_F^m$ depends only on $\mathrm{Tr}_m(F)$ and then evaluating it at the $\mathfrak{O}_F/\mathfrak{p}_F^m$-points. Using the presentation and this isomorphism, one gets an obvious map $\zeta_m : \mathcal{H}(G(F), I_m) \to \mathcal{H}(G(F'), I'_m)$, when the fields F and F' are m-close (also see [Lemaire 2001] for GL_n), which was shown in [Ganapathy 2015] to be an isomorphism of rings. Hence we obtain a bijection

$$\{\text{Iso. classes of irr. ad. representations } (\Pi, V) \text{ of } G(F) \text{ with } \Pi^{I_m} \neq 0\}$$
$$\longleftrightarrow \{\text{Iso. classes of irr. ad. representations } (\Pi', V') \text{ of } G(F') \text{ with } \Pi'^{I'_m} \neq 0\}.$$

When one wants to prove the Kazhdan isomorphism or its variant for general connected reductive groups, one is naturally led to consider parahoric subgroups, study the reduction of the underlying parahoric group schemes mod $\mathfrak{p}_F^m$, and prove that they are determined by truncated data. That is the goal of the present article. Our proof is different from J.K.Yu's approach of proving (1-2) for the Iwahori group scheme of a split p-adic group. We will use the construction of the parahoric group scheme via the Artin–Weil theorem (see [Landvogt 1996]). Let us summarize the main results of this paper.

First, given a split connected reductive group over $\mathbb{Z}$, one can unambiguously work with this group over an arbitrary field after base change. More generally, given a connected reductive group G over F, we first need to make sense of what it means to give a group G' over F' where F' is suitably close to F. Let us first explain how this is done for quasisplit groups. Let (R, Δ) be a based root datum and let $(G_0, T_0, B_0, \{u_\alpha\}_{\alpha\in\Delta})$ be a pinned, split, connected, reductive $\mathbb{Z}$-group with based root datum (R, Δ). We know that the F-isomorphism classes of quasisplit groups G_q that are F-forms of G_0 are parametrized by the pointed cohomology set $H^1(\Gamma_F, \mathrm{Aut}(R, \Delta))$ (see Theorem 3.2). Let $E_{qs}(F, G_0)_m$ be the set of F-isomorphism classes of quasisplit groups G_q that split (and become isomorphic to G_0) over an at most m-ramified extension of F. It is easy to see that this is parametrized by the cohomology set $H^1(\Gamma_F/I_F^m, \mathrm{Aut}(R, \Delta))$. Using the Deligne isomorphism, we prove that there is a bijection $E_{qs}(F, G_0)_m \to E_{qs}(F', G'_0)_m$, $G_q \to G'_q$, provided F and F' are m-close. Moreover, with the cocycles chosen compatibly, this will yield data (G_q, T_q, B_q) over F (where T_q is a maximal F-torus and B_q is an F-Borel containing T_q), and correspondingly (G'_q, T'_q, B'_q) over F', together with an isomorphism $X_*(T_q) \to X_*(T'_q)$ that is Del_m-equivariant (see Lemma 3.4). It is a simple observation that the maximal F-split subtorus S_q of T_q is a maximal F-split torus in G_q (see Lemma 4.1). We prove that there is a simplicial isomorphism between the apartments $\mathcal{A}_m : \mathcal{A}(S_q, F) \to \mathcal{A}(S'_q, F')$ if the fields F and F' are m-close (see Proposition 4.4 for precise statement). Let $\mathcal{F}$ be a facet in $\mathcal{A}(S_q, F)$ and $\mathcal{F}' = \mathcal{A}_m(\mathcal{F})$. Then $\mathcal{F}'$ is a facet in $\mathcal{A}(S'_q, F')$. We prove that the parahoric group schemes $\mathcal{P}_{\mathcal{F}} \times_{\mathfrak{O}_F} \mathfrak{O}_F/\mathfrak{p}_F^m$ and $\mathcal{P}_{\mathcal{F}'} \times_{\mathfrak{O}_{F'}} \mathfrak{O}_{F'}/\mathfrak{p}_{F'}^m$ are isomorphic provided F and F' are l-close for $l \gg m$ (see Theorem 4.5 and Proposition 4.10 for precise statements). To prove this theorem, we prove an analogous statement for the root subgroup schemes if the fields F and F' are sufficiently close, invoke the result of Chai–Yu [2001] that the reduction of the (lft) Néron models of the corresponding tori are isomorphic if the fields are sufficiently close, and use the Artin–Weil theorem on obtaining group schemes as solutions to birational group laws.

To move to the general case, we recall that any connected reductive group is an inner form of a quasisplit group, and the F-isomorphism classes of inner forms of G_q is parametrized by the cohomology set $H^1(\mathrm{Gal}(F_{\mathrm{un}}/F), G_q^{\mathrm{ad}}(F_{\mathrm{un}}))$ (where F_{un} is the maximal unramified extension of F contained in F_s). With G'_q corresponding to G_q as above, we prove in Lemma 5.1 that

$$H^1(\mathrm{Gal}(F_{\mathrm{un}}/F), G_q^{\mathrm{ad}}(F_{\mathrm{un}})) \cong H^1(\mathrm{Gal}(F'_{\mathrm{un}}/F'), G_q^{\mathrm{ad}}(F'_{\mathrm{un}}))$$

as pointed sets if the fields F and F' are m-close using the work of Kottwitz [2014]. Using the work of DeBacker and Reeder [2009] it is further possible to refine the above and obtain an isomorphism at the level of cocycles (see Section 5.A). All the above yields data (G, S, A) where G is a connected reductive group over F that is an inner form of G_q, a maximal F_{un}-split F-torus S that contains a maximal F-split torus A of G, and similarly (G', S', A') over F', together with a $\mathrm{Gal}(\widehat{F_{\mathrm{un}}}/F)$-equivariant simplicial isomorphism $\mathcal{A}_{m,*} : \mathcal{A}(S, \widehat{F_{\mathrm{un}}}) \to \mathcal{A}(S', \widehat{F'_{\mathrm{un}}})$ (see Lemma 6.1). Here $\widehat{F_{\mathrm{un}}}$ denotes the completion of F_{un}. Let $\tilde{\mathcal{F}}_*$ be a $\mathrm{Gal}(\widehat{F_{\mathrm{un}}}/F)$-invariant facet in $\mathcal{A}(S, \widehat{F_{\mathrm{un}}})$ and let $\tilde{\mathcal{F}}'_* = \mathcal{A}_{m,*}(\tilde{\mathcal{F}}_*)$. We prove that there is a $\mathrm{Gal}(\widehat{F_{\mathrm{un}}}/F)$-equivariant isomorphism

$$\tilde{p}_{m,*} : \mathcal{P}_{\tilde{\mathcal{F}}_*} \times_{\mathfrak{O}_{\widehat{F_{\mathrm{un}}}}} \mathfrak{O}_{\widehat{F_{\mathrm{un}}}}/\mathfrak{p}^m_{\widehat{F_{\mathrm{un}}}} \to \mathcal{P}_{\tilde{\mathcal{F}}'_*} \times_{\mathfrak{O}_{\widehat{F'_{\mathrm{un}}}}} \mathfrak{O}_{\widehat{F'_{\mathrm{un}}}}/\mathfrak{p}^m_{\widehat{F'_{\mathrm{un}}}}$$

provided F and F' are l-close (see Proposition 6.2). With $\mathcal{F}_* := (\tilde{\mathcal{F}}_*)^{\mathrm{Gal}(\widehat{F_{\mathrm{un}}}/F)}$ and $\mathcal{F}'_* := (\tilde{\mathcal{F}}'_*)^{\mathrm{Gal}(\widehat{F'_{\mathrm{un}}}/F')}$, the above descends to an isomorphism of group schemes

$$p_{m,*} : \mathcal{P}_{\mathcal{F}_*} \times_{\mathfrak{O}_F} \mathfrak{O}_F/\mathfrak{p}_F^m \to \mathcal{P}_{\mathcal{F}'_*} \times_{\mathfrak{O}_{F'}} \mathfrak{O}_{F'}/\mathfrak{p}_{F'}^m.$$

As a corollary, we obtain that

$$\mathcal{P}_{\mathcal{F}_*}(\mathfrak{O}_F/\mathfrak{p}_F^m) \cong \mathcal{P}_{\mathcal{F}'_*}(\mathfrak{O}_{F'}/\mathfrak{p}_{F'}^m)$$

as groups provided the fields F and F' are l-close.

2. Some review

Unless otherwise stated, F will denote a nonarchimedean local field, that is, a complete discretely valued field with perfect residue field. Let $\mathfrak{O}_F$ denote its ring of integers, $\mathfrak{p}_F$ its maximal ideal, $\omega = \omega_F$ an additive valuation on F normalized so that $\omega(F) = \mathbb{Z}$, and $\pi = \pi_F$ a uniformizer. Fix a separable closure F_s of F and let $\Gamma_F = \mathrm{Gal}(F_s/F)$.

2.A. ***Deligne's theory.*** Let $m \geq 1$. Let I_F be the inertia group of F and I_F^m be its m-th higher ramification subgroup with upper numbering (see Chapter IV of [Serre 1979]). Let us summarize the results of Deligne [1984] that will be used later in this article. Deligne considered the triplet $\mathrm{Tr}_m(F) = (\mathfrak{O}_F/\mathfrak{p}_F^m, \mathfrak{p}_F/\mathfrak{p}_F^{m+1}, \epsilon)$, where ϵ is the natural projection of $\mathfrak{p}_F/\mathfrak{p}_F^{m+1}$ on $\mathfrak{p}_F/\mathfrak{p}_F^m$, and proved that Γ_F/I_F^m is canonically determined by $\mathrm{Tr}_m(F)$. Hence an isomorphism of triplets $\psi_m : \mathrm{Tr}_m(F) \to \mathrm{Tr}_m(F')$ gives rise to an isomorphism

$$\Gamma_F/I_F^m \xrightarrow{\mathrm{Del}_m} \Gamma_{F'}/I_{F'}^m \tag{2-1}$$

that is unique up to inner automorphisms (see Equation 3.5.1 of [Deligne 1984]). More precisely, given an integer $f \geq 0$, let $\mathrm{ext}(F)^f$ denote the category of finite separable extensions E/F satisfying the following condition: The normal closure E_1 of E in F_s satisfies $\mathrm{Gal}(E_1/F)^f = 1$. Deligne proved that an isomorphism $\psi_m : \mathrm{Tr}_m(F) \to \mathrm{Tr}_m(F')$ induces an equivalence of categories $\mathrm{ext}(F)^m \to \mathrm{ext}(F')^m$. Here is a partial description of the map Del_m (see Section 1.3 of [Deligne 1984]). Let L be a finite totally ramified Galois extension of F satisfying $I(L/F)^m = 1$ (here $I(L/F)$ is the inertia group of L/F). Then $L = F(\alpha)$ where α is a root of an Eisenstein polynomial

$$P(x) = x^n + \pi \sum a_i x^i$$

for $a_i \in \mathfrak{O}_F$. Let $a'_i \in \mathfrak{O}_{F'}$ be such that $a_i \mod \mathfrak{p}^m \to a'_i \mod \mathfrak{p}'^m$. So a'_i is well-defined mod $\mathfrak{p}'^m$. Then the corresponding extension L'/F' can be obtained as $L' = F'(\alpha')$ where α' is a root of the polynomial

$$P'(x) = x^n + \pi' \sum a'_i x^i$$

where $\pi \mod \mathfrak{p}_F^m \to \pi' \mod \mathfrak{p}_{F'}^m$. The assumption that $I(L/F)^m = 1$ ensures that the extension L' does not depend on the choice of a'_i, up to a unique isomorphism.

2.B. ***The main theorem of Chai–Yu.*** Let T be a torus over F and let K/F be a Galois extension such that T is split over K. Let $\Gamma_{K/F} = \mathrm{Gal}(K/F)$ and let $\Lambda = X_*(T)$, the cocharacter group of T. Then T is determined by the Γ-module Λ up to a canonical isomorphism. With F' denoting another nonarchimedean local field, we will denote the analogous objects over F' with a superscript $'$. We introduce the following series of congruence notation:

- $(\mathfrak{O}_F, \mathfrak{O}_K) \equiv_{\psi_m} (\mathfrak{O}_{F'}, \mathfrak{O}_{K'})$ (level m): this means that ψ_m is an isomorphism $\mathfrak{O}_K/\pi^m\mathfrak{O}_K \to \mathfrak{O}_{K'}/\pi'^m\mathfrak{O}_{K'}$ and induces an isomorphism $\mathfrak{O}_F/\pi^m\mathfrak{O}_F \to \mathfrak{O}_{F'}/\pi'^m\mathfrak{O}_{F'}$. We denote this induced isomorphism also by ψ_m. Having chosen the uniformizers, this also induces an isomorphism $\mathrm{Tr}_m(F) \to \mathrm{Tr}_m(F')$, which we still denote by ψ_m.
- $(\mathfrak{O}_F, \mathfrak{O}_K, \Gamma_{K/F}) \equiv_{\psi_m,\gamma} (\mathfrak{O}_{F'}, \mathfrak{O}_{K'}, \Gamma_{K'/F'})$(level m): this means $(\mathfrak{O}_F, \mathfrak{O}_K) \equiv_{\psi_m} (\mathfrak{O}_{F'}, \mathfrak{O}_{K'})$(level m), γ is an isomorphism $\Gamma_{K/F} \to \Gamma_{K'/F'}$, and ψ_m is $\Gamma_{K/F}$-equivariant relative to γ.
- $(\mathfrak{O}_F, \mathfrak{O}_K, \Gamma_{K/F}, \Lambda) \equiv_{\psi_m,\gamma,\lambda} (\mathfrak{O}_{F'}, \mathfrak{O}_{K'}, \Gamma_{K'/F'}, \Lambda')$ (level m): this means $(\mathfrak{O}_F, \mathfrak{O}_K, \Gamma_{K/F}) \equiv_{\alpha,\beta} (\mathfrak{O}_{F'}, \mathfrak{O}_{K'}, \Gamma_{K'/F'})$ (level m) and λ is an isomorphism $\Lambda \to \Lambda'$ which is $\Gamma_{K/F}$-equivariant relative to γ.

We say that "X is determined by $(\mathfrak{O}_F/\pi^m\mathfrak{O}_F, \mathfrak{O}_K/\pi^m\mathfrak{O}_K, \Gamma_{K/F}, \Lambda)$″ to mean that if

$$(\mathfrak{O}_F, \mathfrak{O}_K, \Gamma_{K/F}, \Lambda) \equiv_{\psi_m,\gamma,\lambda} (\mathfrak{O}_{F'}, \mathfrak{O}_{K'}, \Gamma_{K'/F'}, \Lambda')(\text{level } m)$$

then there is a canonical $\Gamma_{K/F}$-equivariant isomorphism $X \to X'$ determined by $(\psi_m, \gamma, \lambda)$.

Let $\mathcal{T}^{NR}$ denote the Néron–Raynaud model of T considered in [Chai and Yu 2001]. This is a smooth model of T with connected generic fiber such that $\mathcal{T}^{NR}(\mathfrak{O}_{\widehat{F_{\mathrm{un}}}})$ is the maximal bounded subgroup of $T(\widehat{F_{\mathrm{un}}})$, where $\widehat{F_{\mathrm{un}}}$ is the completion of the maximal unramified extension F_{un} of F contained in F_s. This model is of finite type over $\mathfrak{O}_F$.

Theorem 2.1 [Chai and Yu 2001, Theorem 8.5]. *Let $m \geq 1$. There exists $l \geq m$ such that the model*

$$\mathcal{T}^{NR} \times_{\mathfrak{O}_F} \mathfrak{O}_F/\mathfrak{p}_F^m \quad \textit{is determined by} \quad (\mathfrak{O}_F/\pi^l\mathfrak{O}_F, \mathfrak{O}_K/\pi^l\mathfrak{O}_K, \Gamma_{K/F}, \Lambda).$$

The parameters that l depends on are also explicitly determined in Theorem 8.5 of [Chai and Yu 2001]. Let $\mathcal{T}$ denote the neutral component of $\mathcal{T}^{NR}$. This is a smooth model over $\mathfrak{O}_F$ with connected generic and special fibers, and is of finite type over $\mathfrak{O}_F$. Its $\mathfrak{O}_{\widehat{F_{\text{un}}}}$-points is the Iwahori subgroup of $T(\widehat{F_{\text{un}}})$.

Lemma 2.2. *Let $\mathcal{T}$, $l \geq m$ as above. Then the model*

$$\mathcal{T} \times_{\mathfrak{O}_F} \mathfrak{O}_F/\mathfrak{p}_F^m \quad \textit{is determined by} \quad (\mathfrak{O}_F/\pi^l\mathfrak{O}_F, \mathfrak{O}_K/\pi^l\mathfrak{O}_K, \Gamma_{K/F}, \Lambda).$$

Proof. This lemma follows from Lemma 8.5 of [Chai and Yu 2001] and the observation that the formation of $\mathcal{T}$ commutes with any base change on $\text{Spec}(\mathfrak{O}_F)$, that is,

$$(\mathcal{T}^{NR} \times_{\mathfrak{O}_F} \mathfrak{O}_F/\mathfrak{p}_F^m)^0 = \mathcal{T} \times_{\mathfrak{O}_F} \mathfrak{O}_F/\mathfrak{p}_F^m. \qquad \square$$

When the connected reductive group is a torus T, the model $\mathcal{T}$ is its Iwahori (or parahoric) group scheme. We will study congruences of parahoric group schemes attached to facets in the Bruhat–Tits building of a connected reductive group G over F. To this end, let us recall some results from Bruhat–Tits theory and the construction of parahoric group schemes (using Artin–Weil theorem, following [Landvogt 1996]), that will used later in this article.

Given a connected reductive group G over F, let G^{der} denote the derived subgroup of G, and G^{ad} its adjoint group. Let $\mathcal{B}(G, F)$ denote the reduced Bruhat–Tits building of G over F, that is, the building of G^{der} over F. The building is obtained by gluing together apartments $\mathcal{A}(S, F)$ where S runs over the maximal F-split tori in G. The apartment $\mathcal{A}(S, F)$ is an affine space under $X_*(S^{\text{der}}) \otimes_{\mathbb{Z}} \mathbb{R}$ where $S^{\text{der}} = S \cap G^{\text{der}}$. Let $\mathcal{F}$ be a facet in $\mathcal{B}(G, F)$ and let $P_{\mathcal{F}}$ denote the parahoric subgroup of $G(F)$ attached to $\mathcal{F}$. Bruhat–Tits show that there exists a smooth affine $\mathfrak{O}_F$-group scheme $\mathcal{P}_{\mathcal{F}}$ with generic fiber G such that $\mathcal{P}_{\mathcal{F}}(\mathfrak{O}_F) = P_{\mathcal{F}}$. We recall the construction of $\mathcal{P}_{\mathcal{F}}$, following Landvogt [1996]. The parahoric group scheme is first constructed over $\widehat{F_{\text{un}}}$ (note that $G_{\widehat{F_{\text{un}}}}$ is quasisplit), and the model over F is obtained using étale descent.

2.C. ***Structure of quasisplit groups.*** Let G denote a quasisplit connected reductive group over F. Let S be a maximal F-split torus in G and let T and N be the centralizer and normalizer of S in G, respectively. Let B be an F-Borel subgroup of G with $T \subset B$. Note that T is a maximal F-torus in G. Further G and T split over F_s and the Galois group Γ_F acts on the group of characters $X^*(T)$ of T, preserves the root system $\Phi(G, T)$ of T in G, and also the base $\tilde{\Delta}$ of $\Phi(G, T)$ associated to the Borel subgroup B. Let $K \subset F_s$ denote the smallest subextension of F_s splitting T (and hence G). Let $\Phi(G, S)$ denote the set of roots of S in G.

2.C.1. *Root subgroups U_a, $a \in \Phi(G, S)$.* The elements of $\Phi(G, S)$ are restrictions of elements of $\Phi(G, T)$ to S, and the restrictions to S of the elements of $\tilde{\Delta}$ form a basis Δ of $\Phi(G, S)$. Moreover, the elements of $\tilde{\Delta}$ that have the same restriction to S form a single Galois orbit for the action of Γ_F on $\tilde{\Delta}$.

For $\alpha \in \Phi(G, T)$, let $\tilde{U}_\alpha$ be the corresponding root subgroup of G_K. The group $\Gamma_{K/F}$ permutes $\tilde{U}_\alpha$ and $\gamma(\tilde{U}_\alpha) = \tilde{U}_{\gamma(\alpha)}$. Let Σ_α be the stabilizer of $\tilde{U}_\alpha$ and let L_α be the corresponding field of invariants. We say that L_α is the *field of definition* of α. Note that $\tilde{U}_\alpha$ is defined over L_α by Galois descent. Let $\{\tilde{x}_\alpha : \mathbb{G}_{a,L_\alpha} \to \tilde{U}_\alpha \mid \alpha \in \Phi(G, T)\}$ denote a Chevalley–Steinberg splitting of G. It has the following properties:

(a) If the restriction a of $\alpha \in \Phi(G, T)$ to S is an indivisible element of $\Phi(G, S)$, then $\tilde{x}_\alpha$ is an L_α-isomorphism of $\mathbb{G}_a$ to $\tilde{U}_\alpha$ and we have $\tilde{x}_{\gamma(\alpha)} = \gamma \circ \tilde{x}_\alpha \circ \gamma^{-1}$ for each $\gamma \in \mathrm{Gal}(K/F)$.

(b) If the restriction a of $\alpha \in \Phi(G, T)$ to S is divisible, then there exists two distinct roots $\beta, \beta' \in \Phi(G, T)$ of restriction $a/2$ to S such that $\alpha = \beta + \beta'$; we have $L_\beta = L_{\beta'}$, L_β is a quadratic separable extension of L_α and for each $\gamma \in \mathrm{Gal}(K/F)$ there exists $\epsilon = \pm 1$ such that $\gamma \circ \tilde{x}_\alpha(u) \circ \gamma^{-1} = \tilde{x}_{\gamma(\alpha)}(\epsilon u)$; if $\gamma \in \mathrm{Gal}(K/L_\alpha)$, we have $\epsilon = -1$ if and only if γ induces the unique nontrivial automorphism of L_β.

Now we describe all possible structures for the root subgroups U_a, $a \in \Phi(G, S)$. We may and do assume that $a \in \Delta$. Let $\tilde{\Delta}_a$ be the orbit of $\Gamma_{K/F}$ in $\tilde{\Delta}$. Let $\pi : G^a \to \langle U_a, U_{-a} \rangle$ be the universal cover of the semisimple group generated by U_a and U_{-a}. The classification of Dynkin diagrams gives two possible cases:

Case I. The group G^a_K is isomorphic to a product of the groups SL_2 indexed by $\tilde{\Delta}_a$ and are permuted transitively by $\mathrm{Gal}(K/F)$, the field of definition of the factor of index α is L_α and $G^a \cong \mathrm{Res}_{L_\alpha/F}\, \mathrm{SL}_2$. Then $U_a \cong \mathrm{Res}_{L_\alpha/F} \tilde{U}_\alpha$ for $\alpha \in \tilde{\Delta}_a$. If $\tilde{x}_\alpha : L_\alpha \to \tilde{U}_\alpha$, then $x_a = \mathrm{Res}_{L_\alpha/F}\, \tilde{x}_\alpha$ is a F-isomorphism of $\mathrm{Res}_{L_\alpha/F}\, \mathbb{G}_a$ to U_a; the pair (L_α, x_a) is called a pinning of U_a. Via x_a, we obtain an isomorphism of L_α with $U_a(F)$, which we also denote by x_a. If $(\tilde{x}_\beta)_{\beta \in \tilde{\Delta}}$ is an Chevalley–Steinberg splitting of G, then we have for each $u \in L_\alpha$,

$$x_a(u) = \prod_{\beta \in \tilde{\Delta}_a} \tilde{x}_\beta(u_\beta) \tag{2-2}$$

In the above, $\beta = \gamma(\alpha)$ for some $\gamma \in \Gamma_{K/F}$ and $u_\beta := \gamma(u)$. The subgroups U_{-a} and the splitting x_{-a} are obtained using $U_{-\alpha}$ and $\tilde{x}_{-\alpha}$ analogously.

Case II. The group G^a_K is isomorphic to a product of the groups SL_3 indexed by the set I consisting of pairs of two elements $\{\alpha, \bar{\alpha}\}$ of $\tilde{\Delta}_a$ such that $\alpha + \bar{\alpha}$ is a root. We have $L_\alpha = L_{\bar{\alpha}}$, L_α is a quadratic extension of $L_{\alpha+\bar{\alpha}}$. The simple factor $\overline{G}$ of index $\{\alpha, \bar{\alpha}\}$ is defined over $L_{\alpha+\bar{\alpha}}$, split over L_α, and is isomorphic over $L_{\alpha+\bar{\alpha}}$ to the special unitary group of the Hermitian form $h : (x_{-1}, x_0, x_1) \to \tau(x_{-1})x_1 + \tau(x_0)x_0 + \tau(x_1)x_{-1}$ over L^3. Here τ is the unique nontrivial element of $\mathrm{Gal}(L_\alpha/L_{\alpha+\bar{\alpha}})$. We denote this simple factor as SU_3, and then $G^a \cong \mathrm{Res}_{L_{\alpha+\bar{\alpha}}/F}\, \mathrm{SU}_3$.

Let $H_0(L_\alpha, L_{\alpha+\bar{\alpha}}) := \{(u, v) \in L_\alpha \times L_\alpha \mid v + \tau(v) = u\tau(u)\}$ denote the $L_{\alpha+\bar{\alpha}}$-group with group law $(u, v) \cdot (\tilde{u}, \tilde{v}) = (u + \tilde{u}, v + \tilde{v} + \tau(u)\tilde{u})$. Then $\zeta : (u, v) \to \tilde{x}_\alpha(u)\tilde{x}_{\alpha+\bar{\alpha}}(-v)\tilde{x}_{\bar{\alpha}}(\tau(u))$ is an $L_{\alpha+\bar{\alpha}}$-group isomorphism of $H_0(L_\alpha, L_{\alpha+\bar{\alpha}})$ with the subgroup $\overline{U} = \tilde{U}_\alpha \tilde{U}_{\alpha+\bar{\alpha}} \tilde{U}_{\bar{\alpha}}$ of $\overline{G}$. Then $U_a = \mathrm{Res}_{L_{\alpha+\bar{\alpha}}/F}\, \overline{U}$ and $x_a = \mathrm{Res}_{L_{\alpha+\bar{\alpha}}/K}\, \zeta$ is an F-isomorphism of groups $H(L_\alpha, L_{\alpha+\bar{\alpha}}) = \mathrm{Res}_{L_{\alpha+\bar{\alpha}}/F}\, H_0(L_\alpha, L_{\alpha+\bar{\alpha}})$ with U_a. Further, for $(u, v) \in L_\alpha \times L_\alpha$,

$$x_a(u, v) = \prod \tilde{x}_\beta(u_\beta)\tilde{x}_{\beta+\bar{\beta}}(-v_\beta)\tilde{x}_{\bar{\beta}}(\tau(u_\beta))$$

In the above, for each β, we choose $\gamma \in \mathrm{Gal}(K/F)$ such that $\beta = \gamma(\alpha)$; then $\bar{\beta} = \gamma(\bar{\alpha})$, $\tilde{x}_\beta = \gamma \circ \tilde{x}_\alpha \circ \gamma^{-1}$, $\tilde{x}_{\bar{\beta}} = \gamma \circ \tilde{x}_{\bar{\alpha}} \circ \gamma^{-1}$, $\tilde{x}_{\beta+\bar{\beta}} = \gamma \circ \tilde{x}_{\alpha+\bar{\alpha}} \circ \gamma^{-1}$, $u_\beta = \gamma(u)$, $v_\beta = \gamma(v)$.

Note that the root subgroup $U_{2a}(K)$ associated to the root $2a$ consists of elements $x_a(0, v)$ where v is an element of $L^0_\alpha := \{v \in L_\alpha \mid v + \tau(v) = 0\}$, and the map $v \to x_a(0, v)$ is an F-vector space isomorphism of L^0_α.

2.C.2. *On the splitting extension of the root.* Let $a \in \Phi^{\mathrm{red}}(G, S)$ with $2a$ is not a root. We fix a pinning (L_α, x_a) of U_a where $\alpha \in \tilde{\Delta}_a$ as in Case I above. The subset of endomorphisms of the F-vector space U_a of the form $\mu_{x_a}(t) : x_a(u) \to x_a(tu)$ for $t \in L_\alpha$ does not depend on the choice of (L_α, x_a) (see Section 4.1.8 of [Bruhat and Tits 1984]). This is denoted by L_a and is called the field attached of the root a. It is isomorphic to L_α via the map $t \to \mu_{x_a}(t)$. Its inverse gives an embedding of $L_a \hookrightarrow K$. A similar definition is obtained when $2a$ is a root in Section 4.1.14 of [Bruhat and Tits 1984].

2.C.3. *Valuations.* Let $\omega : F \to \mathbb{R}^\times$ be as in Section 2.A, and we denote its extension to K also as ω. The notion of valuation of root datum was defined in [Bruhat and Tits 1972]. For $\alpha \in \Phi(G, T)$, put

$$\phi_\alpha(\tilde{x}_\alpha(u)) = \omega(u), \quad u \in K^\times.$$

Then $\tilde{\phi} = (\phi_\alpha)_{\alpha\in\Phi(G,T)}$ defines a valuation of the root datum $(T_K, (\tilde{U}_\alpha)_{\alpha\in\Phi(G,T)})$ in the group $G(K)$ (recall that G_K is split). It is shown in [Bruhat and Tits 1984] that $\tilde{\phi}$ descends to $(T, (\tilde{U}_a)_{a\in\Phi(G,S)})$ and defines a valuation on it. We explicitly define $\phi_a : U_a(F)\backslash\{1\} \to \mathbb{R}$ from $\tilde{\phi}$. For $a \in \Phi(G, S)$, let A (resp. B) be the set of $\alpha \in \Phi(G, T)$ whose restriction to S is a (resp. $2a$). For $u \in U_a(F)$, there exist unique $\tilde{u}_\alpha$ such that $u = \prod_{\alpha\in A\cup B} \tilde{u}_\alpha$ for an arbitrary ordering of $A \cup B$ and we put

$$\phi_a(u) = \inf\Big(\inf_{\alpha\in A} \tilde{\phi}_\alpha(\tilde{u}_\alpha), \inf_{\alpha\in B} \tfrac{1}{2}\tilde{\phi}_\alpha(\tilde{u}_\alpha)\Big).$$

This number is independent of the choice of ordering of $A \cup B$. Then $\phi = (\phi_a)_{a\in\Phi(G,S)}$ defines a valuation of root datum on $(T, (U_a)_{a\in\Phi(G,S)})$ (see Section 4.2.2 of [Bruhat and Tits 1984]).

2.D. ***Parahoric group schemes: quasisplit descent.*** In this section, we assume that F is also strictly Henselian, that is its residue field is separably closed.

2.D.1. *Affine root system and the associated Weyl groups.* The apartment $\mathcal{A}(S, F)$ can also be thought of as the set of valuations that are equipollent to $\phi = (\phi_a)_{a\in\Phi(G,S)}$, where ϕ as above. This is an affine space under $X_*(S^{\mathrm{der}}) \otimes_{\mathbb{Z}} \mathbb{R}$ and $N(F)$ acts on it by affine transformations (see Section 6.2.2 of [Bruhat and Tits 1972]). Let us denote the point of $\mathcal{A}(S, F)$ corresponding to ϕ as x_0. For $a \in \Phi(G, S)$, let $\Gamma_a = \phi_a(U_a(F)\backslash\{1\})$ and

$$\tilde{\Gamma}_a = \{\phi_a(u) \mid u \in U_a(F)\backslash\{1\}, \phi_a(u) = \sup \phi_a(uU_{2a}(F))\}.$$

Here we have used the convention that $U_{2a} = 1$ if $2a$ is not a root. Let

$$\Phi^{\mathrm{af}}(G, S) = \{\psi : \mathcal{A}(S, F) \to \mathbb{R} \mid \psi(\cdot) = a(\cdot - x_0) + l, a \in \Phi(G, S), l \in \tilde{\Gamma}_a\}$$

denote the set of affine roots of S in G. Choosing x_0 allows us to identify $A(S,F)$ with $X_*(S^{\mathrm{der}})\otimes_{\mathbb{Z}}\mathbb{R}$. With this identification, the vanishing hyperplanes coming from $\Phi(G,S)^{\mathrm{af}}$ makes $\mathcal{A}(S,F)$ into a (poly)simplicial complex. The group generated by reflections through the hyperplanes coming from $\Phi(G,S)^{\mathrm{af}}$ is the affine Weyl group denoted by W^{af}. The extended affine Weyl group is defined as $W^e := N(F)/T(F)_1$ where $T(F)_1$ is the kernel of the Kottwitz homomorphism $\kappa_T : T(F)\to X^*(\hat{T}^{I_F}) = X_*(T)_{I_F}$ (see [Haines and Rapoport 2008]). With $W := W(G,S)$, the group W^e hence fits into an exact sequence

$$1\to X_*(T)_{I_F}\to W^e\to W\to 1.$$

2.D.2. *The associated root subgroup schemes.* Let us recall the filtrations on root subgroups and the associated root subgroup schemes from Section 4.3 of [Bruhat and Tits 1984]. For $a\in\Phi(G,S)$, let $\phi_a : U_a(F)\to\mathbb{R}\cup\{\infty\}$ be as above. For $k\in\mathbb{R}$, Let $U_{a,k}=\{u\in U_a(F)\mid \phi_a(u)\geq k\}$. Next, let us describe the associated root subgroup schemes.

Case I. Let $a\in\Phi^{\mathrm{red}}(G,S)$ such that $2a\notin\Phi(G,S)$. For $k\in\tilde{\Gamma}_a$, let $L_{a,k}=\{u\in L_a\mid\omega(u)\geq k\}$. Then $L_{a,k}$ is a free $\mathfrak{O}_F$-module of finite type. Let $\mathcal{L}_{a,k}$ be the canonical smooth $\mathfrak{O}_F$-group scheme associated to this module. (More precisely, given a free $\mathfrak{O}_F$-module M of finite type, the functor taking any $\mathfrak{O}_F$-algebra R to the additive group $R\otimes M$ is representable by a smooth $\mathfrak{O}_F$-group scheme $\mathcal{M}$ whose affine algebra is identified with the symmetric algebra of the dual of M.) Let $U_{a,k}$ be the image under x_a of $L_{a,k}$ and let $\mathcal{U}_{a,k}$ be the $\mathfrak{O}_F$-group scheme obtained by transport of structure using x_a. Then $\mathcal{U}_{a,k}$ has generic fiber U_a and $\mathcal{U}_{a,k}(\mathfrak{O}_F)=U_{a,k}$. The definition is extended to $k\in\mathbb{R}\backslash\{0\}$ in Section 4.3.2 of [Bruhat and Tits 1984].

Case II. Let $a\in\Phi^{\mathrm{red}}(G,S)$ with $2a\in\Phi(G,S)$. The root subgroup $U_a\cong\mathrm{Res}_F^{L_{2a}} H_0(L_a,L_{2a})$ via x_a. In order to describe the root subgroup schemes of the filtration $U_{a,k}$, we use an alternate description of $H_0(L_a,L_{2a})$. Recall that L_a^0 is the set of trace 0 elements of L_a. Let L_a^1 denote the set of trace 1 elements in L_a and let

$$(L_a)^1_{\max} := \{\lambda\in L_a^1\mid\omega(\lambda)=\sup\{\omega(x)\mid x\in L_a^1\}\}.$$

Note that $(L_a)^1_{\max}\neq\varnothing$ and when the residue field of L_a is of characteristic $\neq 2$, $1/2\in(L_a)^1_{\max}$. Let $\lambda\in(L_a)^1_{\max}$ and let $H_0^\lambda := L_a\times L_a^0$ equipped with the action

$$(u,v)\cdot(\tilde{u},\tilde{v}) = (u+\tilde{u},\, v+\tilde{v}-\lambda u\tau(\tilde{u})+\tau(\lambda)\tau(u)\tilde{u}). \tag{2-3}$$

Then H_0^λ is an algebraic L_{2a}-group and $j_\lambda : (u,v)\to(u,v-\lambda\tau(u)u)$ is an L_{2a}-group isomorphism of $H_0(L_a,L_{2a})$ onto H_0^λ. Let $H^\lambda=\mathrm{Res}_F^{L_{2a}} H_0^\lambda$.

Let $\gamma=-\frac{1}{2}\omega(\lambda)$. For $k\in\tilde{\Gamma}_a$, let $l=2k+\frac{1}{e_a}$, and

$$L_{a,k+\gamma} := \{u\in L_a\mid\omega(u)\geq k+\gamma\}\quad\text{and}\quad L^0_{a,l} := \{u\in L_a^0\mid\omega(u)\geq l\}.$$

Up to isomorphism, there exists a unique smooth affine $\mathfrak{O}_F$-group scheme $\mathcal{H}_k^\lambda$ of finite type with generic fiber H^λ and such that $\mathcal{H}_k^\lambda(\mathfrak{O}_F)=L_{a,k+\gamma}\times L^0_{a,l}$ and a group law, which induces the group law (2-3) on the generic fiber (See Section 4.3.5 of [Bruhat and Tits 1984]). In more detail, let $\mathcal{L}_{a,k+\gamma}$ and $\mathcal{L}^0_{a,l}$ be the canonical $\mathfrak{O}_{L_{2a}}$-group schemes associated to $L_{a,k+\gamma}$ and $L^0_{a,l}$. Let $\mathcal{H}^\lambda_{0,k} = \mathcal{L}_{a,k+\gamma} \times \mathcal{L}^0_{a,l}$. The map

$$L_a \times L_a \to L^0_a, (u, u') \to \lambda u \tau(\tilde{u}) - \tau(\lambda)\tau(u)\tilde{u}$$

can be extended uniquely to a morphism $\mathcal{L}_{a,k+\gamma} \times \mathcal{L}_{a,k+\gamma} \to \mathcal{L}^0_{a,l}$. Hence the group law can be extended to $\mathcal{H}^\lambda_{0,k}$. Let $\mathcal{H}^\lambda_k := \mathrm{Res}^{\mathfrak{O}_{L_{2a}}}_F H^\lambda_{0,k}$. By transport of structure using $x_a \circ \mathrm{Res}^{L_{2a}}_F j^{-1}_\lambda$, we obtain the $\mathfrak{O}_F$-group scheme $\mathcal{U}_{a,k}$. These definitions are extended to $k, l \in \mathbb{R}\backslash\{0\}$ in Section 4.3.8 of [Bruhat and Tits 1984].

Using the isomorphism $v \to x_a(0, v)$ from $L^0_a \to U_{2a}$, we obtain from the scheme $\mathcal{L}^0_k$ (for $k \in \omega(L^0_a)\backslash 0$), an $\mathfrak{O}_F$-scheme whose generic fiber is U_{2a} and denote it as $\mathcal{U}_{2a,k}$ (see Section 4.3.7 of [Bruhat and Tits 1984] for further details).

2.D.3. *Construction of parahoric group schemes over F.* In this section, we recall the construction of parahoric group schemes, following [Landvogt 1996]. Given $x \in \mathcal{A}(S, F)$, let $f_x : \Phi(G, S) \to \mathbb{R}$ be the function $f_x(a) = -a(x - x_0)$, where x_0 is the unique point arising from quasisplit descent as in Section 2.D.1. Let $U_{a,x} := U_{a, f_x(a)}$. Let $\mathcal{U}_{a,x}$ be the smooth affine group scheme over $\mathfrak{O}_F$ with generic fiber U_a and with $\mathcal{U}_{a,x}(\mathfrak{O}_F) = U_{a,x}$ (as in Section 2.D.2). For $\Psi = \Phi^+(G, S)$ and $\Psi = \Phi^-(G, S)$, Proposition 3.3.2 of [Bruhat and Tits 1984] gives a unique smooth affine $\mathfrak{O}_F$-group scheme $\mathcal{U}_{\Psi,x}$ of finite type with generic fiber U_Ψ and the property that for every good ordering of Ψ^{red} (See Section 3.1.2 of [Bruhat and Tits 1984]), the F-isomorphism $\prod_{a\in\Psi} U_a \to U_\Psi$ can be extended to an $\mathfrak{O}_F$-isomorphism $\prod_{a\in\Psi} \mathcal{U}_{a,x} \to \mathcal{U}_{\Psi,x}$.

The parahoric subgroup P_x is generated by $\mathcal{T}(\mathfrak{O}_F)$ and the $U_{a,x}$ for $a \in \Phi(G, S)$ (with $\mathcal{T}$ is as in Section 2.B). One of the main results of [Bruhat and Tits 1984] is that there is a unique smooth affine $\mathfrak{O}_F$-group scheme $\mathcal{P}_x$ with generic fiber G and with $\mathcal{P}_x(\mathfrak{O}_F) = P_x$. We recall the construction of $\mathcal{P}_x$ from [Landvogt 1996]. The idea is to put an $\mathfrak{O}_F$-birational group law on $\mathcal{U}_{\Phi^+,x} \times \mathcal{T} \times \mathcal{U}_{\Phi^-,x}$ and invoke Artin–Weil theorem (see Chapters 5 and 6 of [Bosch et al. 1990]) to construct $\mathcal{P}_x$. Let us first introduce some notation. Let $\mathcal{U}^\pm_x = \mathcal{U}_{\Phi^\pm(G,S),x}$ and let $\mathcal{X}_x = \mathcal{U}^-_x \mathcal{T} \mathcal{U}^+_x$. Since its generic fiber $\mathcal{X}_x \times_{\mathfrak{O}_F} F = U^- T U^+$ is an open neighborhood of the 1-section of G, there exists a unique F-birational group law on the generic fiber of $\mathcal{X}_x$. We want to extend this to $\mathcal{X}_x$. Since U^-TU^+ and U^+TU^- are both open neighborhoods of the 1-section of G, there exist $f \in F[U^-TU^+]$ and $f' \in F[U^+TU^-]$ such that $F[U^-TU^+]_f = F[U^+TU^-]_{f'}$. Without loss of generality, we may assume that $f \in \mathfrak{O}_F[U^-TU^+]\backslash\pi\mathfrak{O}_F[U^-TU^+]$ and $f' \in \mathfrak{O}_F[U^+TU^-]\backslash\pi\mathfrak{O}_F[U^+TU^-]$. Proposition 5.16 of [Landvogt 1996] shows that inside $F[U^-TU^+]_f = F[U^+TU^-]_{f'}$, we have $\mathfrak{O}_F[\mathcal{U}^-_x\mathcal{T}\mathcal{U}^+_x]_f = \mathfrak{O}_F[\mathcal{U}^+_x\mathcal{T}\mathcal{U}^-_x]_{f'}$. So we will identify $(\mathcal{U}^-_x\mathcal{T}\mathcal{U}^+_x)_f = (\mathcal{U}^+_x\mathcal{T}\mathcal{U}^-_x)_{f'}$ in the following. By Proposition 5.8 of [Landvogt 1996], we can identify $\mathcal{T}\mathcal{U}^+_x$ and $\mathcal{U}^+_x\mathcal{T}$ and hence also $\mathcal{T}\mathcal{U}^+_x\mathcal{U}^-_x$ and $\mathcal{U}^+_x\mathcal{T}\mathcal{U}^-_x$. In $\mathcal{X}_x \times \mathcal{X}_x = \mathcal{U}^-_x \times (\mathcal{T} \times \mathcal{U}^+_x \times \mathcal{U}^-_x) \times \mathcal{T}\mathcal{U}^+_x$, we consider the open subscheme

$$\begin{aligned}\mathcal{U}^-_x \times (\mathcal{U}^+_x \times \mathcal{T} \times \mathcal{U}^-_x)_f \times \mathcal{T}\mathcal{U}^+_x &= \mathcal{U}^-_x \times (\mathcal{U}^-_x \times \mathcal{T} \times \mathcal{U}^+_x)_{f'} \times \mathcal{T}\mathcal{U}^+_x \\ &\subset \mathcal{U}^-_x \times \mathcal{U}^-_x \times \mathcal{T} \times \mathcal{U}^+_x \times \mathcal{T}\mathcal{U}^+_x \\ &= (\mathcal{U}^-_x \times \mathcal{U}^-_x) \times (\mathcal{T} \times \mathcal{T}) \times (\mathcal{U}^+_x \times \mathcal{U}^+_x) \\ &\xrightarrow{\text{mult}^3} \mathcal{U}^-_x \times \mathcal{T} \times \mathcal{U}^+_x\end{aligned}$$

So we obtain a morphism $\mathcal{U}_x^- \times (\mathcal{U}_x^+ \times \mathcal{T} \times \mathcal{U}_x^-)_f \times \mathcal{T}\mathcal{U}_x^+ \to \mathcal{X}_x$. Since $\mathcal{X}_x$ has irreducible fibers over $\mathfrak{O}_F$ and since $f \notin \pi\mathfrak{O}_F[\mathcal{U}_x^- \mathcal{T}\mathcal{U}_x^+]$, we see that $(\mathcal{U}_x^- \mathcal{T}\mathcal{U}_x^+)_f$ is $\mathfrak{O}_F$-dense in $\mathcal{X}_x$ (that is, each of its fibers is Zariski dense in the corresponding fiber of $\mathcal{X}_x$ — see Section 2.5 of [Bosch et al. 1990]), and hence $\mathcal{U}_x^- \times (\mathcal{U}_x^+ \times \mathcal{T} \times \mathcal{U}_x^-)_f \times \mathcal{T}\mathcal{U}_x^+$ is $\mathfrak{O}_F$-dense in $\mathcal{U}_x^- \times (\mathcal{T} \times \mathcal{U}_x^+ \times \mathcal{U}_x^-) \times \mathcal{T}\mathcal{U}_x^+ = \mathcal{X}_x \times \mathcal{X}_x$. Hence we obtain an $\mathfrak{O}_F$-rational map $m : \mathcal{X}_x \times \mathcal{X}_x \to \mathcal{X}_x$. By Proposition 5.16 of [Landvogt 1996], m is an $\mathfrak{O}_F$-birational group law on $\mathcal{X}_x$. Glue together the schemes G and $\mathcal{X}_x$ along $\mathcal{X}_x \times_{\mathfrak{O}_F} F$ and denote it as $\mathcal{Y}_x$. As in Proposition 5.17 of [Landvogt 1996], the parahoric group scheme $\mathcal{P}_x$ with group law $\bar{m}$, together with an open immersion $\mathcal{Y}_x \to \mathcal{P}_x$ such that the restriction of $\bar{m}$ to $\mathcal{Y}_x$ is m, is obtained by applying Theorem 5.1 of [Bosch et al. 1990] to the scheme $\mathcal{Y}_x$. The generic fiber of $\mathcal{P}_x$ is G. Let $\mathcal{F}$ be a facet in $\mathcal{A}(S, F)$. Then for $x, y \in \mathcal{F}$, $P_x = P_y$. So we write $P_{\mathcal{F}}$ for the parahoric subgroup attached to the facet $\mathcal{F}$ and denote the underlying group scheme as $\mathcal{P}_{\mathcal{F}}$.

2.E. *Parahoric group schemes: étale descent.* Let F be a nonarchimedean local field and $\widehat{F_{\mathrm{un}}}$ be the completion of the maximal unramified extension $F_{\mathrm{un}} (\subset F_s)$ of F. Let G be a connected reductive group over F. By a theorem of Steinberg (recalled as Theorem 5.2), we know that $G_{\widehat{F_{\mathrm{un}}}}$ is quasisplit. Let A be a maximal F-split torus in G. By Section 5 of [Bruhat and Tits 1984], there is an F-torus S that contains A and is maximal F_{un}-split. Note that $X_*(A) = X_*(S)^{\mathrm{Gal}(\widehat{F_{\mathrm{un}}}/F)}$. Let $\mathcal{A}(A, F)$ denote the apartment of G with respect to A. Let $\mathcal{F}_*$ be a facet in $\mathcal{A}(A, F)$. We fix an algebraic closure $\bar{\kappa}_F$ of the residue field κ_F and identify the Galois groups $\mathrm{Gal}(\widehat{F_{\mathrm{un}}}/F)$ with $\mathrm{Gal}(\bar{\kappa}_F/\kappa_F)$. Let σ denote the Frobenius element of $\mathrm{Gal}(F_{\mathrm{un}}/F)$ under this identification. Then we know that there is a σ-stable facet $\tilde{\mathcal{F}}_*$ in $\mathcal{A}(S, F_{\mathrm{un}})$ such that $\tilde{\mathcal{F}}_*^{\sigma} = \mathcal{F}_*$ (see Chapter 5 of [Bruhat and Tits 1984]). Since $\tilde{\mathcal{F}}_*$ is stable under the action of σ, the parahoric group scheme $\mathcal{P}_{\tilde{\mathcal{F}}_*}$ is also stable under the action of σ. In this case, the $\mathfrak{O}_{\widehat{F_{\mathrm{un}}}}$-group scheme $\mathcal{P}_{\tilde{\mathcal{F}}_*}$ admits a unique descent to an $\mathfrak{O}_F$-group scheme with generic fiber G (see Example B, Section 6.2, [Bosch et al. 1990]). The affine ring of this group scheme is $(\mathfrak{O}_{\widehat{F_{\mathrm{un}}}}[\mathcal{P}_{\tilde{\mathcal{F}}_*}])^{\mathrm{Gal}(\widehat{F_{\mathrm{un}}}/F)}$. This is the parahoric group scheme attached to the facet $\mathcal{F}_*$ of $\mathcal{A}(A, F)$.

3. Quasisplit forms over close local fields

Let G_0 be a split connected reductive group defined over $\mathbb{Z}$ with root datum (R, Δ). For an extension K/F, let $G_{0,K} := G_0 \times_{\mathbb{Z}} K$.

Let $E(F, G_0)$ be the of F-isomorphism classes of connected reductive F-algebraic groups G with G_{F_s} isomorphic to G_{0,F_s}. This is in natural bijection with the Galois cohomology set $H^1(\Gamma_F, \mathrm{Aut}(G_{0,F_s}))$. We denote this map

$$E(F, G_0) \to H^1(\Gamma_F, \mathrm{Aut}(G_{0,F_s})), [G] \to s_G. \tag{3-1}$$

Lemma 3.1. *Let I_F be the inertia group of F and I_F^m denote the m-th higher ramification subgroup with upper numbering. Let $E(F, G_0)_m$ denote the set of F-isomorphism classes of F-forms G of $G_{0,F}$ such that there exists an at most m-ramified finite extension $L \subset F_s$ (i.e.,* $\mathrm{Gal}(L/F)^m = 1$*) with*

$G \times_F L \cong G_0 \times_{\mathbb{Z}} L$. The bijection (3-1) induces a bijection between $E(F, G_0)_m$ and the cohomology set $H^1(\Gamma_F/I_F^m, (\mathrm{Aut}_{F_s}(G_{0,F_s}))^{I_F^m})$.

Proof. Let $\Omega := (F_s)^{I_F^m}$. Then for every finite extension $F \subset L \subset F_s$, $L \hookrightarrow \Omega$ if and only if $\mathrm{Gal}(L/F)^m = 1$ (see Section 3.5 of [Deligne 1984]). Further we know that $H^1(\mathrm{Aut}(\Omega/F), \mathrm{Aut}_\Omega(G_{0,\Omega}))$ classifies isomorphism classes of F-forms $[G]$ with $G \times_F \Omega \cong G_{0,F} \times_F \Omega$. Now simply note that $\mathrm{Aut}(\Omega/F) \cong \Gamma_F/I_F^m$ and $\mathrm{Aut}_\Omega(G_{0,\Omega}) = (\mathrm{Aut}_{F_s}(G_{0,F_s}))^{I_F^m}$. □

3.A. *Quasisplit forms.* Let $(G_0, T_0, B_0, \{u_\alpha\}_{\alpha\in\tilde{\Delta}})$ be a pinned, split, connected, reductive $\mathbb{Z}$-group with based root datum $(R, \tilde{\Delta})$ where $\{u_\alpha\}_{\alpha\in\tilde{\Delta}}$ is a splitting as in Section 3.2.2 of [Bruhat and Tits 1984]. Then $\mathrm{Out}(G_0)$ can be identified with the constant $\mathbb{Z}$-group scheme associated to the group $\mathrm{Aut}(R, \tilde{\Delta})$. Consider the exact sequence

$$1 \to \mathrm{Inn}(G_0(F_s)) \to \mathrm{Aut}(G_{0,F_s}) \to \mathrm{Aut}(R, \tilde{\Delta}) \to 1.$$

Let $H = H(G_0, T_0, B_0, \{u_\alpha\}_{\alpha\in\tilde{\Delta}})$ be the subgroup of $\mathrm{Aut}(G_{0,F_s})$ consisting of all a such that $a(B_0) = B_0$, $a(T_0) = T_0$ and $\{a \circ u_\alpha \mid \alpha \in \tilde{\Delta}\} = \{u_\alpha \mid \alpha \in \tilde{\Delta}\}$. Then $H \hookrightarrow \mathrm{Aut}(G_{0,F_s}) \to \mathrm{Aut}(R, \tilde{\Delta})$ is an isomorphism and $\mathrm{Aut}(G_{0,F_s}) \cong H \ltimes \mathrm{Inn}(G_0(F_s))$. Hence the natural map $H^1(\Gamma_F, \mathrm{Aut}(G_{0,F_s})) \to H^1(\Gamma_F, \mathrm{Aut}(R, \tilde{\Delta}))$ has a section given by

$$q : H^1(\Gamma_F, \mathrm{Aut}(R, \tilde{\Delta})) \xrightarrow{\sim} H^1(\Gamma_F, H) \to H^1(\Gamma_F, \mathrm{Aut}(G_{0,F_s})).$$

We now recall the following well-known theorem (see [Conrad 2011], Section 7.2).

Theorem 3.2. *Let $[G] \in E(F, G_0)$. Then s_G lies in the image of*

$$q : H^1(\Gamma_F, \mathrm{Aut}(R, \tilde{\Delta})) \to H^1(\Gamma_F, \mathrm{Aut}(G_{0,F_s}))$$

if and only if G is quasisplit over F, that is, it has a Borel subgroup defined over F.

Let $E_{qs}(F, G_0) := \{[G] \in E(F, G_0) \mid s_G \in Im(q)\}$ and $E_{qs}(F, G_0)_m = E_{qs}(F, G_0) \cap E(F, G_0)_m$. Since G_0 is F-split, the action of Γ_F on (G_0, B_0, T_0) is trivial. Hence

$$Z^1(\Gamma_F, \mathrm{Aut}(R, \tilde{\Delta})) = \mathrm{Hom}(\Gamma_F, \mathrm{Aut}(R, \tilde{\Delta})).$$

Lemma 3.3. *We have the following*:

(a) *The class $[G] \in E_{qs}(F, G_0)_m$ if and only if s_G lies in the image of*

$$q : H^1(\Gamma_F/I_F^m, \mathrm{Aut}(R, \tilde{\Delta})^{I_F^m}) \to H^1(\Gamma_F/I_F^m, \mathrm{Aut}(G_{0,F_s})^{I_F^m}).$$

(b) *The isomorphism $\psi_m : \mathrm{Tr}_m(F) \xrightarrow{\sim} \mathrm{Tr}_m(F')$ induces an isomorphism*

$$\mathfrak{Q}_m : H^1(\Gamma_F/I_F^m, \mathrm{Aut}(R, \tilde{\Delta})) \xrightarrow{\sim} H^1(\Gamma_{F'}/I_{F'}^m, \mathrm{Aut}(R, \tilde{\Delta}))$$

and

$$\mathfrak{Q}_m^c : Z^1(\Gamma_F/I_F^m, \mathrm{Aut}(R, \tilde{\Delta})) \xrightarrow{\sim} Z^1(\Gamma_{F'}/I_{F'}^m, \mathrm{Aut}(R, \tilde{\Delta})).$$

(c) *The isomorphism ψ_m induces a bijection $E_{qs}(F, G_0)_m \to E_{qs}(F', G_0)_m$, $[G] \to [G']$, where $s_{G'} = q' \circ \mathfrak{Q}_m(s_G)$.*

Proof. This is clear from Lemma 3.1 and Theorem 3.2. □

As noted in Lemma 3.3, $Z^1(\mathrm{Gal}(\Omega/F), \mathrm{Aut}(R, \tilde{\Delta})) = \mathrm{Hom}(\mathrm{Gal}(\Omega/F), \mathrm{Aut}(R, \tilde{\Delta}))$ since G_0 is split. Let us fix $s \in Z^1(\mathrm{Gal}(\Omega/F), \mathrm{Aut}(R, \tilde{\Delta})) \cong Z^1(\mathrm{Gal}(\Omega/F), H)$. Let (G, ϕ) be a pair of be a quasisplit connected reductive group over F and $\phi : G_0 \times_{\mathbb{Z}} \Omega \to G \times_F \Omega$ an Ω-isomorphism such that the Galois action on $G(F_s)$ is given by s. We may and do assume that there is a finite Galois at most m-ramified extension K of F over which ϕ is defined, that is, that $s \in Z^1(\mathrm{Gal}(K/F), \mathrm{Aut}(R, \tilde{\Delta}))$.

More precisely, with $*_F$ denoting the Galois action on $G(K)$, we have

$$\gamma *_F \phi(x) = \phi(s(\gamma)(\gamma \cdot x))$$

for $\gamma \in \mathrm{Gal}(K/F)$ and $x \in G_0(K)$. Then $\phi(T_0) = T$ is a maximal torus of G defined over F and $\phi(B_0) = B$ is a Borel subgroup of G containing T and defined over F. Let $s' \in Z^1(\mathrm{Gal}(K'/F'), \mathrm{Aut}(R, \tilde{\Delta}))$ as in Lemma 3.3. Here K'/F' is determined by K/F via Del_m. Let (G', ϕ') be a pair of quasisplit connected reductive group over F' and $\phi' : G_0 \times_{\mathbb{Z}} K' \to G' \times_{F'} K'$ such that $\gamma' *_{F'} \phi'(x') = \phi(s'(\gamma')(\gamma' \cdot x'))$, where $\gamma' = \mathrm{Del}_m(\gamma)$. Then $\phi'(T_0) = T'$ and $\phi'(B_0) = B'$ are defined over F'. Note that $X_*(T) \cong X_*(T_0) \cong X_*(T')$ and $X^*(T) \cong X^*(T_0) \cong X^*(T')$ via ϕ and ϕ'.

Recall the notation of Chai–Yu: $(\mathfrak{O}_F, \mathfrak{O}_K, \Gamma_{K/F}) \equiv_{\psi_m, \gamma} (\mathfrak{O}_{F'}, \mathfrak{O}_{K'}, \Gamma_{K'/F'})$ (level m) from Section 2.B. We write

$$(\mathfrak{O}_F, \mathfrak{O}_K, \Gamma_{K/F}, H) \equiv_{\psi_m, \gamma, \mathfrak{Q}_m^c} (\mathfrak{O}_{F'}, \mathfrak{O}'_{K'}, \Gamma_{K'/F'}, H')(\text{level } m)$$

to mean $(\mathfrak{O}_F, \mathfrak{O}_K, \Gamma_{K/F}) \equiv_{\alpha,\beta} (\mathfrak{O}_{F'}, \mathfrak{O}'_{K'}, \Gamma_{K'/F'})$ (level m), H and H' arise from the same $\mathbb{Z}$-pinned group $(G_0, B_0, T_0, \{u_\alpha\}_{\alpha \in \tilde{\Delta}})$, and the F-quasisplit data (G, B, T) with cocycle s corresponds to the F'-quasisplit data (G', B', T') with cocycle s' via $\mathfrak{Q}_m^c$ as in Lemma 3.3 (b) (but applied to K and K' respectively). To abbreviate notation we will write *congruence data* D_m to mean

$$D_m : (\mathfrak{O}, \mathfrak{O}_K, \Gamma_{K/F}, H) \equiv_{\psi_m, \gamma, \mathfrak{Q}_m^c} (\mathfrak{O}_{F'}, \mathfrak{O}'_{K'}, \Gamma_{K'/F'}, H')(\text{level } m).$$

Lemma 3.4. *The congruence data D_m induces isomorphisms*:

$$X^*(T)^{\mathrm{Gal}(\Omega/F)} \cong X^*(T')^{\mathrm{Gal}(\Omega'/F')}, \quad X_*(T)^{\mathrm{Gal}(\Omega/F)} \cong X_*(T')^{\mathrm{Gal}(\Omega'/F')},$$
$$X^*(T)_{\mathrm{Gal}(\Omega/F)} \cong X^*(T')_{\mathrm{Gal}(\Omega'/F')}, \quad X_*(T)_{\mathrm{Gal}(\Omega/F)} \cong X_*(T')_{\mathrm{Gal}(\Omega'/F')}.$$

Proof. We know that $\gamma *_F (\phi(x)) = \phi(s(\gamma)(\gamma \cdot x))$ where $s(\gamma) = \phi^{-1} \circ \gamma(\phi)$ takes values in $H = H(G_0, T_0, B_0, \{u_\alpha\}_{\alpha \in \tilde{\Delta}})$. We similarly have $*_{F'}$. This action induces the action on $X_*(T)$ as follows:

$$\gamma *_F (\phi \circ \lambda) = \phi \circ (s(\gamma)(\lambda))$$

where $\gamma \in \mathrm{Gal}(\Omega/F)$ and $\lambda \in X_*(T_0)$, where we now view $s(\gamma)$ as an element of $\mathrm{Aut}(R, \tilde{\Delta})$. By definition $s(\gamma)(\lambda) = s'(\gamma')(\lambda)$ where $\gamma' = \mathrm{Del}_m(\gamma)$. Hence $\gamma' *_{F'} (\phi' \circ \lambda) = \phi' \circ s'(\gamma')(\lambda)$. Now, $X_*(T)^{\mathrm{Gal}(\Omega/F)} = \{\phi \circ \lambda \mid s(\gamma)(\lambda) = \lambda\}$. The lemma is now clear. □

4. Congruences of parahoric group schemes: quasisplit descent

4.A. ***Apartment over close local fields.*** In this section, we additionally assume that F is strictly Henselian. We begin with the following lemma.

Lemma 4.1. *Let T as above and let S be the maximal split subtorus of T. Then S is maximal F-split and $Z_G(S) = T$.*

Proof. Let $S \subset \tilde{S}$ with $\tilde{S}$ maximal F-split. Since G is quasisplit over F, $\tilde{T} = Z_G(\tilde{S})$ is a maximal torus in G and we can assume that $\tilde{T} \subset \tilde{B}$, with $\tilde{B}$ defined over F. Then B and $\tilde{B}$ are $G(F)$-conjugate, which implies that T and $\tilde{T}$ are $G(F)$-conjugate. But conjugation by an element of $G(F)$ will preserve the split and anisotropic components of T, which implies that S and $\tilde{S}$ are $G(F)$-conjugate, which forces $S = \tilde{S}$ to be maximal F-split. It is now clear that $Z_G(S) = T$. □

Remark 4.2. The torus $S^{\text{der}} := S \cap G^{\text{der}}$ is a maximal F-split torus of G^{der} contained in $T^{\text{der}} := T \cap G^{\text{der}}$.

4.A.1. *Compatibility of Chevalley–Steinberg systems.* Recall that we have fixed a $\mathbb{Z}$-pinning $\{u_\alpha\}_{\alpha\in\Delta}$ of G_0. This, via the Galois action given by the cocycles s and s', gives rise to a Steinberg splitting $\{x_\alpha\}_{\alpha\in\Delta}$ of G and a Steinberg splitting $\{x'_{\alpha'}\}_{\alpha'\in\Delta'}$ of G' respectively. Let $\Phi_m : \Phi(G, T) \xrightarrow{\sim} \Phi(G', T')$ (since both are isomorphic to $\Phi(G_0, T_0)$). This isomorphism is Del_m-equivariant. Note that with $\gamma \in \text{Gal}(\Omega/F)$ and $\gamma' = \text{Del}_m(\gamma)$, we have that $x_{\gamma(\alpha)} = \gamma \circ x_\alpha \circ \gamma^{-1}$ and $x'_{\gamma'(\alpha')} = \gamma' \circ x'_{\alpha'} \circ \gamma'^{-1}$ where $\alpha' = \Phi_m(\alpha)$. The $\{x_\alpha\}_{\alpha\in\Delta}$ and $\{x'_{\alpha'}\}_{\alpha'\in\Delta'}$ each extend to Chevalley–Steinberg systems on G and G' respectively and continue to have the compatibility with Del_m in the sense described above.

We define

$$e_F := \begin{cases} e_{F/\mathbb{Q}_2} = \omega_F(2) & \text{if } \operatorname{char}(F) = 0 \text{ and residue } \operatorname{char}(F) = 2, \\ \infty & \text{otherwise.} \end{cases}$$

We prove the following refinement of Lemma 4.3.3 of [Bruhat and Tits 1984] when the residue characteristic of F is 2, using the additional hypothesis that the extension K/F splitting G is at most m-ramified.

Lemma 4.3. *Let $m \geq 1$ and let F be of residue characteristic 2 with $e_F \geq m$. Let G, B, T as above, where G splits over K with $\text{Gal}(K/F)^m = 1$. Assume that $a, 2a \in \Phi(G, S)$. Consider the separable quadratic extension L_a/L_{2a} inside K. Let $e_a = e_{L_a/F}$, $e_{2a} = e_{L_{2a}/F}$. There exists $t \in L_a$ with $L_a = L_{2a}[t]$ and the coefficients $A, B \in L_{2a}$ of the equation $t^2 + At + B = 0$ satisfied by t have the following properties*:

(a) *$\omega(B) = 0$ or B is a uniformizer of L_{2a}.*

(b) *$\omega(B) \leq \omega(A) < \frac{m}{2} + \frac{1}{e_a}$.*

In particular $A \neq 0$.

Proof. By Lemma 4.3.3(ii) of [Bruhat and Tits 1984], (a) holds, and $A = 0$ or $\omega(B) \leq \omega(A) < \omega(2)$ or $0 < \omega(B) \leq \omega(A) = \omega(2)$. Since $\text{Gal}(K/F)^m = \text{Gal}(K/F)_{\psi_{K/F}(m)} = 1$ where $\psi_{K/F}$ denotes the inverse of the Herbrand function (See Chapter 4 of [Serre 1979]), we have

$$\text{Gal}(K/L_{2a})^{\psi_{L_{2a}/F}(m)} = \text{Gal}(K/L_{2a})_{\psi_{K/F}(m)} = \text{Gal}(K/L_{2a}) \cap \text{Gal}(K/F)_{\psi_{K/F}(m)} = 1.$$

This implies that $\mathrm{Gal}(L_a/L_{2a})^{\psi_{L_{2a}/F}(m)} = 1$. Using the equivalence of (ii) and (iv) of Lemma A.6.1 of [Deligne 1984], we see that

$$\omega(\tau(t) - t) < \frac{\psi_{L_{2a}/F}(m) + 1}{2e_{2a}} = \frac{\psi_{L_{2a}/F}(m) + 1}{e_a}. \tag{4-1}$$

It is easy to see from the definition that $\psi_{L_{2a}/F}(m) \leq m \cdot e_{2a}$. Hence

$$\omega(\tau(t) - t) < \tfrac{m}{2} + \tfrac{1}{e_a}.$$

Now, $\omega(A) = \omega(\tau(t) + t) \geq \min(\omega(\tau(t) - t), \omega(2t))$, and $\omega(2t) = \omega(2) + \omega(t) = e_F + \frac{1}{e_a}$. Since $e_F \geq m > m/2$, we see that

$$\omega(A) = \min(\omega(\tau(t) - t), \omega(2t)) = \omega(\tau(t) - t) < \tfrac{m}{2} + \tfrac{1}{e_a} \tag{4-2}$$

and in particular, $A \neq 0$.

Note that when the characteristic of F is 2, the claim that $A \neq 0$ simply follows from the fact that the extension L_a/L_{2a} is separable. □

Proposition 4.4. *Let G, T and B as in the preceding paragraph. Let $m \geq 1$ and let F, F' be such that $e_F, e_{F'} \geq m$. The congruence data D_m induces a simplicial isomorphism $\mathcal{A}_m : \mathcal{A}(S, F) \to \mathcal{A}(S', F')$, where (G', B', T') corresponds to the triple (G, B, T) as above and S (resp. S') is the maximal split subtorus of T (resp. T') which is maximal F-split (resp. F'-split) by Lemma 4.1. Furthermore, with W^e as in Section 2.D.1, we also have a group isomorphism $W^e \cong W^{e'}$.*

Proof. The reduced apartment $\mathcal{A}(S, F)$ is an affine space under $X_*(S^{\mathrm{der}}) \otimes_{\mathbb{Z}} \mathbb{R}$. Using Lemma 3.4, we see that D_m induces a unique bijection $\mathcal{A}_m : \mathcal{A}(S, F) \to \mathcal{A}(S', F')$ such that $x_0 \to x_0'$ (where x_0, x_0' are as in Section 2.D.1 arising from Chevalley–Steinberg systems chosen compatibly as in Section 4.A.1).

It remains to observe that $\mathcal{A}_m$ is a simplicial isomorphism. Recall that the elements of $\Phi(G, S)$ are restrictions to S of the elements of $\Phi(G, T)$ and two elements of $\Phi(G, T)$ restrict to the same element of $\Phi(G, S)$ if and only if they lie in the same $\mathrm{Gal}(K/F)$-orbit. Further, with $\tilde{\Delta}$ denoting a base of $\Phi(G, T)$, the elements $\alpha|_S, \alpha \in \tilde{\Delta}$ form a base Δ of $\Phi(G, S)$. Let $\Phi_m : \Phi(G, T) \xrightarrow{\sim} \Phi(G', T')$ (since both are isomorphic to $\Phi(G_0, T_0)$). This isomorphism is Del_m-equivariant. Hence the obvious map $\Phi(G, S) \to \Phi(G', S')$, $\alpha|_S \to \Phi_m(\alpha)|_{S'}$, which we also denote as Φ_m, is an isomorphism of the relative root systems. (In more detail, since S and S' have the same rank, we have a isomorphism of $\mathbb{R}$-vector spaces $X^*(S) \otimes_{\mathbb{Z}} \mathbb{R} \to X^*(S') \otimes_{\mathbb{Z}} \mathbb{R}$. Further, we have a bijection between $\Delta \to \Delta'$; this is because if $\Phi_m(\alpha)|_{S'} = \Phi_m(\beta)|_{S'}$, then there is $\eta' \in \mathrm{Gal}(\Omega'/F')$ with $\eta' \cdot \Phi_m(\alpha) = \Phi_m(\beta)$. Then $\eta \cdot \alpha = \beta$ where $\eta' = \mathrm{Del}_m(\eta)$. Finally note that $\langle \Phi_m(\alpha)|_{S'}, \Phi_m(\beta)|_{S'} \rangle = \langle \Phi_m(\alpha), \Phi_m(\beta) \rangle = \langle \alpha, \beta \rangle = \langle \alpha|_S, \beta|_S \rangle$).

The vanishing hyperplanes with respect to the affine roots $\Phi^{\mathrm{af}}(G, S)$ gives the simplicial structure on $\mathcal{A}(S, F)$. Recall that

$$\Phi^{\mathrm{af}}(G, F) = \{\psi : \mathcal{A}(S, F) \to \mathbb{R} \mid \psi(\cdot) = a(\cdot - x_0) + l, a \in \Phi(G, S), l \in \tilde{\Gamma}_a\}.$$

For any $a \in \Phi(G, S)$, let $a' = \Phi_m(a)$. Let $L_{a'} \subset K'$ denote splitting extension of the root a' obtained by Del_m. Since F is strictly Henselian, the extensions L_a/F and $L_{a'}/F'$ are totally ramified. To prove that the bijection Φ_m extends to a bijection $\Phi_m^{\mathrm{af}} : \Phi^{\mathrm{af}}(G, F) \to \Phi^{\mathrm{af}}(G', F')$ making $\mathcal{A}_m$ a simplicial isomorphism, we simply have to observe that for each $a \in \Phi(G, S)$, $\tilde{\Gamma}_a = \tilde{\Gamma}_{a'}$. By Section 4.3.4 of [Bruhat and Tits 1984], we have the following:

Case I. Suppose $a \in \Phi^{\mathrm{red}}(G, S)$, $2a \notin \Phi(G, S)$. Then $\Gamma_a = \tilde{\Gamma}_a = \frac{1}{e_a}\mathbb{Z}$.

Case II. Suppose $a, 2a \in \Phi(G, S)$.

(a) Suppose L_a/L_{2a} is ramified and the residue characteristic of F is not 2. Then

$$\tilde{\Gamma}_a = \tfrac{1}{e_a}\mathbb{Z} \quad \text{and} \quad \tilde{\Gamma}_{2a} = \tfrac{1}{e_a} + \tfrac{1}{e_{2a}}\mathbb{Z}.$$

(b) Suppose L_a/L_{2a} is ramified and the residue characteristic of F is 2. By Lemma 4.3, $A \neq 0$. Then

$$\tilde{\Gamma}_a = \tfrac{1}{2e_a} + \tfrac{1}{e_a}\mathbb{Z} \quad \text{and} \quad \tilde{\Gamma}_{2a} = \tfrac{1}{e_{2a}}\mathbb{Z}.$$

Since $e_a = e_{a'}$, $e_{2a} = e_{2a'}$, and the valuations ω and ω' are normalized so that $\omega(F) = \omega'(F') = \mathbb{Z}$, we have $\tilde{\Gamma}_a = \tilde{\Gamma}_{a'}$ for all $a \in \Phi(G, S)$. □

4.B. *Congruences of parahoric group schemes: strictly Henselian case.* In this section, we additionally assume that F is strictly Henselian.

Theorem 4.5. *Let $m \geq 1$ and let F and F' be such that $e_F, e_{F'} \geq 2m$. Let l be as in Lemma 2.2 and let D_l and G, S, T, B as in the beginning of this section. Let $\mathcal{F} \in \mathcal{A}(S, F)$ and $\mathcal{F}' = \mathcal{A}_m(\mathcal{F})$ as in Proposition 4.4. Let $\mathcal{P}_{\mathcal{F}}$ be the parahoric group scheme over $\mathfrak{O}_F$ attached to $\mathcal{F}$ by Bruhat–Tits, and let $\mathcal{P}_{\mathcal{F}'}$ be the group scheme attached to $\mathcal{F}'$ over $\mathfrak{O}_{F'}$. Then the congruence data D_l induces an isomorphism of group schemes*

$$\tilde{p}_m : \mathcal{P}_{\mathcal{F}} \times_{\mathfrak{O}_F} \mathfrak{O}_F/\mathfrak{p}_F^m \to \mathcal{P}_{\mathcal{F}'} \times_{\mathfrak{O}_{F'}} \mathfrak{O}_{F'}/\mathfrak{p}_{F'}^m \times_{\psi_m^{-1}} \mathfrak{O}_F/\mathfrak{p}_F^m.$$

In particular, $\mathcal{P}_{\mathcal{F}}(\mathfrak{O}_F/\mathfrak{p}_F^m) \cong \mathcal{P}_{\mathcal{F}'}(\mathfrak{O}_{F'}/\mathfrak{p}_{F'}^m)$ as groups.

To prove this theorem, we will study the reduction of root subgroup schemes mod $\mathfrak{p}_F^m$ and prove that they are determined by congruence data, use the result of Chai–Yu that the reduction of the Néron model of the torus in determined by congruence data, study the reduction of $\mathfrak{O}_F$-birational group laws in Section 2.D.3, and invoke the Artin–Weil theorem to obtain the corresponding result for parahoric group schemes in Section 4.B.1.

The following lemma is easy.

Lemma 4.6. *Let M be a free $\mathfrak{O}_F$-module of finite type and let $A = \mathrm{Sym}_{\mathfrak{O}_F}(M^\vee)$ be the symmetric algebra of $M^\vee$, where $M^\vee := Hom_{\mathfrak{O}_F}(M, \mathfrak{O}_F)$. Then*

$$A \otimes_{\mathfrak{O}_F} \mathfrak{O}_F/\mathfrak{p}_F^m \cong \mathrm{Sym}_{\mathfrak{O}_F/\mathfrak{p}_F^m}(M^\vee \otimes_{\mathfrak{O}_F} \mathfrak{O}_F/\mathfrak{p}_F^m) \cong \mathrm{Sym}_{\mathfrak{O}_F/\mathfrak{p}_F^m}(\mathrm{Hom}_{\mathfrak{O}_F/\mathfrak{p}_F^m}(M \otimes_{\mathfrak{O}_F} \mathfrak{O}_F/\mathfrak{p}_F^m, \mathfrak{O}_F/\mathfrak{p}_F^m)).$$

Lemma 4.7. *Let $m \geq 1$, let F and F' be such that $e_F, e_{F'} \geq 2m$ and let D_m as before. Let $a \in \Phi(G, S)$ and $k \in \mathbb{R}$. Let $\mathcal{U}_{a,k}$ be the $\mathfrak{O}_F$-group scheme in Section 2.D.2. Let $a' = \Phi_m(a) \in \Phi(G', S')$ and let $\mathcal{U}'_{a',k}$*

be the $\mathfrak{O}_{F'}$-group scheme in Section 2.D.2. Then the congruence data D_m induces an isomorphism of group schemes

$$\mathcal{U}_{a,k} \times_{\mathfrak{O}_F} \mathfrak{O}_F/\mathfrak{p}_F^m \cong \mathcal{U}_{a',k} \times_{\mathfrak{O}_{F'}} \mathfrak{O}_{F'}/\mathfrak{p}_{F'}^m \times_{\psi_m^{-1}} \mathfrak{O}_F/\mathfrak{p}_F^m.$$

In particular,

$$\mathcal{U}_{a,k}(\mathfrak{O}_F/\mathfrak{p}_F^m) \cong \mathcal{U}_{a',k}(\mathfrak{O}_{F'}/\mathfrak{p}_{F'}^m).$$

Proof. We will stick to the notation in Section 2.D.2.

Case I. Suppose $a \in \Phi^{\text{red}}(G,S)$, $2a \notin \Phi(G,S)$. The affine ring representing $\mathcal{U}_{a,k}$ is isomorphic to $\mathrm{Sym}_{\mathfrak{O}_F} L_{a,k}^\vee$. Note that $L_{a,k} = \mathfrak{p}_{L_a}^{\lceil k/e \rceil}$. Since $\mathfrak{p}_{L_a}$ is a free $\mathfrak{O}_F$-module of rank equal to $[L_a : F]$, it is clear that the data D_m induces an isomorphism of $L_{a,k} \otimes_{\mathfrak{O}_F} \mathfrak{O}_F/\mathfrak{p}_F^m$ and $L_{a',k} \otimes_{\mathfrak{O}_{F'}} \mathfrak{O}_{F'}/\mathfrak{p}_{F'}^m$ and we are done by the previous lemma.

Case II. Suppose $a, 2a \in \Phi(G,S)$. Since F is strictly Henselian, the extension L_a/L_{2a} is totally ramified. Let $L_a = L_{2a}(t)$, where $t^2 + At + B = 0$ with A, B satisfying Lemma 4.3.3 of [Bruhat and Tits 1984]. When:

- The residue characteristic of F is not 2, we take $\lambda = \frac{1}{2}$ (See Lemma 4.3.3(ii) of [loc. cit.]).
- The residue characteristic of F is 2, we take $\lambda = tA^{-1}$ (using Lemma 4.3.3(ii) of [loc. cit.] and Lemma 4.3).

Then the affine ring representing the scheme $\mathcal{H}_0^\lambda$ is

$$\mathrm{Sym}_{\mathfrak{O}_{L_{2a}}} L_{a,k+\gamma}^\vee \otimes_{\mathfrak{O}_{L_{2a}}} \mathrm{Sym}_{\mathfrak{O}_{L_{2a}}} (L_{a,l}^0)^\vee \cong \mathrm{Sym}_{\mathfrak{O}_{L_{2a}}} ((L_{a,k+\gamma} \times L_{a,l}^0)^\vee),$$

where $l = 2k + \frac{1}{e_a}$. We describe $L_{a,l}^0$.

(a) If the residue characteristic of F is not 2, then using that $\omega(2) = 0$ in Lemma 4.3.3 of [loc. cit.], we see that $A = 0$. Then $L_a^0 = \{x \in L_a \mid \tau(x) + x = 0\} = \{yt \mid y \in L_{2a}\}$ and

$$L_{a,l}^0 = \{yt \mid y \in L_{2a}, \omega(yt) \geq l\} = \{yt \mid y \in L_{2a}, \omega(y) \geq 2k\}.$$

(b) If the residue characteristic of F is 2, then:
   (i) If $\mathrm{char}(F) = 2$, then $L_a^0 = L_{2a}$ and $L_{a,l}^0 = \{y \in L_{2a} \mid \omega(y) \geq l\}$.
   (ii) If $\mathrm{char}(F) = 0$, then $L_a^0 = \{y(1 - 2tA^{-1}) \mid y \in L_{2a}\}$. By Lemma 4.3, we have

$$\omega(2tA^{-1}) = e_F + \frac{1}{e_a} - \omega(A) > e_F - \frac{m}{2} \geq m$$

since $e_F \geq 2m$. Hence $1 - 2tA^{-1} \in 1 + \mathfrak{p}_{L_a}^{me_a}$, and $L_{a,l}^0 = \{y(1 - 2tA^{-1}) \mid y \in L_{2a}, \omega(y) \geq l\}$.

Let $L_{a'} \subset \Omega'$ be obtained from L_a via the Deligne isomorphism Del_m. Then $L_{a'}$ is the splitting extension of the root a' (and similarly we obtain $L_{2a'}$). We may and do assume that $L_{a'} = L_{2a'}(t')$, where $t'^2 + A't' + B' = 0$, with A', B' satisfying:

- $\omega(A) = \omega'(A')$ and $A \bmod \mathfrak{p}_{L_{2a}}^{me_{2a}} \xrightarrow{\psi_m} A' \bmod \mathfrak{p}_{L'_{2a'}}^{me_{2a}}$.
- $\omega(B) = \omega'(B')$ and $B \bmod \mathfrak{p}_{L_{2a}}^{me_{2a}} \xrightarrow{\psi_m} B' \bmod \mathfrak{p}_{L'_{2a'}}^{me_{2a}}$.

Then $t \mod \mathfrak{p}_{L_a}^{me_a} \xrightarrow{\psi_m} t' \mod \mathfrak{p}_{L'_{a'}}^{me_a}$. It is now easy to check that the map ψ_m induces isomorphisms

$$L_{a,k+\gamma} \otimes_{\mathfrak{O}_{L_{2a}}} \mathfrak{O}_{L_{2a}}/\mathfrak{p}_{L_{2a}}^{me_{2a}} \cong L^0_{a',k+\gamma} \otimes_{\mathfrak{O}_{L'_{2a'}}} \mathfrak{O}_{L'_{2a'}}/\mathfrak{p}_{L'_{2a'}}^{me_{2a}}$$
$$L^0_{a,l} \otimes_{\mathfrak{O}_{L_{2a}}} \mathfrak{O}_{L_{2a}}/\mathfrak{p}_{L_{2a}}^{me_{2a}} \cong L^0_{a',l} \otimes_{\mathfrak{O}_{L'_{2a'}}} \mathfrak{O}_{L'_{2a'}}/\mathfrak{p}_{L'_{2a'}}^{me_{2a}}.$$

In the above, we have used that when the residue characteristic of F is 2, $1 - 2tA^{-1} \equiv 1 \mod \mathfrak{p}_{L_a}^{me_a}$. Consequently, D_m induces an isomorphism of the reduction of the respective affine rings mod $\mathfrak{p}_{L_{2a}}^{me_{2a}}$. To see that this is an isomorphism of group schemes, we observe that reducing the map

$$\begin{aligned} j : L_{a,k} \times L^0_{a,l} \times L_{a,k} \times L^0_{a,l} &\to L_{a,k} \times L^0_{a,l} \\ ((x, y), (x', y')) &\to (x + x', y + y' - \lambda x \tau(x') + \lambda x' \tau(x)) \end{aligned}$$

mod $\mathfrak{p}_{L_{2a}}^{me_{2a}}$ is ψ_m-equivariant. Finally $\mathcal{H}^\lambda = \operatorname{Res}_{\mathfrak{O}_F}^{\mathfrak{O}_{L_{2a}}} \mathcal{H}_0^\lambda$ and the result now follows from [Bosch et al. 1990, page 192].

The lemma for $\mathcal{U}_{2a,k}$ follows using that

$$L^0_{a,k} \otimes_{\mathfrak{O}_{L_{2a}}} \mathfrak{O}_{L_{2a}}/\mathfrak{p}_{L_{2a}}^{me_{2a}} \cong L^0_{a',k} \otimes_{\mathfrak{O}_{L'_{2a'}}} \mathfrak{O}_{L'_{2a'}}/\mathfrak{p}_{L'_{2a'}}^{me_{2a}}$$

and [Bosch et al. 1990, page 192]. □

The following corollary is an obvious consequence of the previous lemma.

Corollary 4.8. *With assumptions of Lemma 4.7, and with $\mathcal{F}' = \mathcal{A}_m(\mathcal{F})$ where $\mathcal{F}$ is a facet in $\mathcal{A}(S, F)$, let $\mathcal{U}_{a,\mathcal{F}}$ (resp. $\mathcal{U}_{a',\mathcal{F}'}$) be the smooth root subgroup scheme over $\mathfrak{O}_F$ (resp. $\mathfrak{O}_{F'}$) as in Section 2.D.3. The congruence data D_m induces an isomorphism*

$$\mathcal{U}_{a,\mathcal{F}} \times_{\mathfrak{O}_F} \mathfrak{O}_F/\mathfrak{p}_F^m \cong \mathcal{U}_{a,\mathcal{F}'} \times_{\mathfrak{O}_{F'}} \mathfrak{O}_{F'}/\mathfrak{p}_{F'}^m \times_{\psi_m^{-1}} \mathfrak{O}_F/\mathfrak{p}_F^m.$$

In particular, $\mathcal{U}_{a,\mathcal{F}}(\mathfrak{O}_F/\mathfrak{p}_F^m) \cong \mathcal{U}_{a',\mathcal{F}'}(\mathfrak{O}_{F'}/\mathfrak{p}_{F'}^m)$ as groups.

4.B.1. *Proof of Theorem 4.5.* For a scheme X defined over a local ring R with maximal ideal $\mathfrak{m}$, we will denote $X^{(m)} := X \times_R R/\mathfrak{m}^m$. Let l be as in Lemma 2.2. We want to prove that D_l induces an isomorphism of $\mathfrak{O}_F/\mathfrak{p}_F^m$-group schemes $\mathcal{P}_{\mathcal{F}}^{(m)} \cong \mathcal{P}_{\mathcal{F}'}^{(m)} \times_{\psi_m^{-1}} \mathfrak{O}_F/\mathfrak{p}_F^m$. Let $\mathcal{X}_{\mathcal{F}}, \mathcal{X}_{\mathcal{F}'}$ be as in Section 2.D.3. Let $m^{(m)}$ be the $\mathfrak{O}_F/\mathfrak{p}_F^m$-birational group law on $\mathcal{X}_{\mathcal{F}}^{(m)}$ and similarly $n^{(m)}$ on $\mathcal{X}_{\mathcal{F}'}^{(m)}$. Note that via D_l, we also have that

$$(\mathfrak{O}_F, \mathfrak{O}_K, \Gamma_{K/F}, \Lambda) \equiv_{\psi_e,\gamma,\lambda} (\mathfrak{O}_{F'}, \mathfrak{O}_{K'}, \Gamma_{K'/F'}, \Lambda') \text{ (level } l)$$

as in the notation of Chai–Yu of Section 2.B, where $\Lambda = X_*(T)$, $\Lambda' = X_*(T')$; so the result of Lemma 2.2 holds. We know by Lemmas 2.2 and 4.8 that

$$\mathcal{X}_{\mathcal{F}}^{(m)} \cong \mathcal{X}_{\mathcal{F}'}^{(m)} \times_{\psi_m^{-1}} \mathfrak{O}_F/\mathfrak{p}_F^m \tag{4-3}$$

as $\mathfrak{O}_F/\mathfrak{p}_F^m$-schemes. Further, by these lemmas, we also have that the $\mathfrak{O}_F/\mathfrak{p}_F^m$-birational group laws $n^{(m)} \times_{\psi_m^{-1}} \mathfrak{O}_F/\mathfrak{p}_F^m$ and $m^{(m)}$ on $\mathcal{X}_{\mathcal{F}}^{(m)}$ are equivalent. Since $\mathcal{Y}_{\mathcal{F}}$ is the $\mathfrak{O}_F$-scheme obtained by gluing G and $\mathcal{X}_{\mathcal{F}}$ along $\mathcal{X}_{\mathcal{F}} \times_{\mathfrak{O}_F} F$, we have that $\mathcal{Y}_{\mathcal{F}}^{(m)}$ is isomorphic to $\mathcal{X}_{\mathcal{F}}^{(m)}$ as $\mathfrak{O}_F/\mathfrak{p}_F^m$-schemes. Now, $\mathcal{P}_{\mathcal{F}}^{(m)}$ with group law $\overline{m}^{(m)}$, and $\mathcal{P}_{\mathcal{F}'}^{(m)} \times_{\psi_m^{-1}} \mathfrak{O}_F/\mathfrak{p}_F^m$ with group law $\bar{n}^{(m)} \times_{\psi_m^{-1}} \mathfrak{O}_F/\mathfrak{p}_F^m$, are both smooth, separated $\mathfrak{O}_F/\mathfrak{p}_F^m$- group schemes that are faithfully flat and of finite type. Recall that the restriction of $\overline{m}$ to $\mathcal{Y}_{\mathcal{F}}$ is m, and similarly for $\bar{n}$. Hence the group laws $\overline{m}^{(m)}$ and $\bar{n}^{(m)} \times_{\psi_m^{-1}} \mathfrak{O}_F/\mathfrak{p}_F^m$ have the same restriction to $\mathcal{Y}_{\mathcal{F}}^{(m)}$. Following the proof of uniqueness of Artin–Weil theorem (see Proposition 3, Section 5.1 of [Bosch et al. 1990]), we obtain that the group schemes $\mathcal{P}_{\mathcal{F}}^{(m)}$ and $\mathcal{P}_{\mathcal{F}'}^{(m)} \times_{\psi_m^{-1}} \mathfrak{O}_F/\mathfrak{p}_F^m$ are isomorphic. □

4.C. ***Congruences of parahoric group schemes: descending from $G_{\widehat{F_{\mathrm{un}}}}$ to G_F.*** In this section, F denotes a nonarchimedean local field and $\widehat{F_{\mathrm{un}}}$ denotes the completion of the maximal unramified extension F_{un} of F. Let A be a maximal F-split torus in G, S maximal F_{un}-split F-torus that contains A. Let $T = Z_G(S)$. Note that $X_*(S) = X_*(T)^{\mathrm{Gal}(\Omega/F_{\mathrm{un}})}$ and $X_*(A) = X_*(T)^{\mathrm{Gal}(\Omega/F)}$.

Lemma 4.9. *The simplicial isomorphism*

$$\mathcal{A}_m : \mathcal{A}(S, \widehat{F_{\mathrm{un}}}) \to \mathcal{A}(S', \widehat{F'_{\mathrm{un}}})$$

of Proposition 4.4 is Del_m*-equivariant.*

Proof. This is clear from the proof of Proposition 4.4, Section 4.A.1, and Lemma 3.4. □

Let $\sigma \in \mathrm{Gal}(\widehat{F_{\mathrm{un}}}/F)$ be as in Section 2.E. Let $\mathcal{F}$ be a facet in $X_*(A)$. Then $\mathcal{F}$ corresponds to a σ-stable facet $\tilde{\mathcal{F}}$ in $X_*(S)$. Note that Del_m induces isomorphisms

$$\mathrm{Gal}(\widehat{F_{\mathrm{un}}}/F) \cong \mathrm{Gal}(F_s/F)/I_F \cong \mathrm{Gal}(F_s'/F')/I_{F'} \cong \mathrm{Gal}(\widehat{F'_{\mathrm{un}}}/F').$$

Let $\sigma' = \mathrm{Del}_m(\sigma)$ under this isomorphism. Let $\tilde{\mathcal{F}}' = \mathcal{A}_m(\tilde{\mathcal{F}})$ and $\mathcal{F}' = \tilde{\mathcal{F}'}^{\sigma'}$.

Proposition 4.10. *The isomorphism*

$$\tilde{p}_m : \mathcal{P}_{\tilde{\mathcal{F}}} \times_{\mathfrak{O}_{\widehat{F_{\mathrm{un}}}}} \mathfrak{O}_{\widehat{F_{\mathrm{un}}}}/\mathfrak{p}^m_{\widehat{F_{\mathrm{un}}}} \to \mathcal{P}_{\tilde{\mathcal{F}}'} \times_{\mathfrak{O}_{\widehat{F'_{\mathrm{un}}}}} \mathfrak{O}_{\widehat{F'_{\mathrm{un}}}}/\mathfrak{p}^m_{\widehat{F'_{\mathrm{un}}}}$$

has the property that $\sigma' \circ \tilde{p}_m = \tilde{p}_m \circ \sigma$.

Proof. Recall that the cocycle s_G has been chosen to take values in $\mathrm{Aut}(H)$ and $s_G \to s_{G'}$ via Lemma 3.3. Further, $\mathcal{T}$ is defined over $\mathfrak{O}_F$ and $\mathcal{T}_{\mathfrak{O}_{\widehat{F_{\mathrm{un}}}}} = \mathcal{T} \times_{\mathfrak{O}_F} \mathfrak{O}_{\widehat{F_{\mathrm{un}}}}$. From this it is clear that $\sigma' \circ \tilde{p}_m = \tilde{p}_m \circ \sigma$ on $\mathcal{T} \times_{\mathfrak{O}_{\widehat{F_{\mathrm{un}}}}} \mathfrak{O}_{\widehat{F_{\mathrm{un}}}}/\mathfrak{p}^m_{\widehat{F_{\mathrm{un}}}}$. In addition, using the fact that Chevalley–Steinberg systems on G and G' have been chosen compatibly (see Section 4.A.1), it is easy to see that $\sigma' \circ \tilde{p}_m = \tilde{p}_m \circ \sigma$ on $\mathcal{U}_{\tilde{\mathcal{F}}} \times_{\mathfrak{O}_{\widehat{F_{\mathrm{un}}}}} \mathfrak{O}_{\widehat{F_{\mathrm{un}}}}/\mathfrak{p}^m_{\widehat{F_{\mathrm{un}}}}$. This completes the proof of the proposition. □

5. Inner forms of quasisplit groups over close local fields

Let F be a nonarchimedean local field and let G be a connected reductive group over F. Then there is a quasisplit group G_q defined over F such that G is an inner form of G_q. In particular, the F-isomorphism class of G is determined by an element in $H^1(\Gamma_F, G_q^{\mathrm{ad}}(F_s))$. Moreover if $[G] \in E(F, G_0)_m$ then $[G_q] \in E(F, G_0)_m$ and $[G]$ is determined by an element of $H^1(\mathrm{Aut}(\Omega/F), G_q^{\mathrm{ad}}(\Omega))$ (Recall that $\Omega = (F_s)^{I_F^m}$). Let s_{G_q} be the element of $H^1(\Gamma_F/I_F^m, \mathrm{Aut}(R, \Delta)^{I_F^m})$ that determines (G_q, B_q, T_q), up to

F-isomorphisms. Let G_q^{der} be the derived subgroup of G_q and let G_q^{ad}, G_q^{sc} denote the corresponding adjoint and simply connected groups. Then the groups G_q^{der}, G_q^{ad}, G_q^{sc} are quasisplit (if S_q is a maximal F-split torus in G_q whose centralizer T_q is a maximal torus, then $S_q \cap G_q^{\mathrm{der}}$ is a maximal F-split torus of G_q^{der} and $Z_{G_q^{\mathrm{der}}}(S_q \cap G_q^{\mathrm{der}}) = T_q \cap G_q^{\mathrm{der}}$ is a maximal torus of G_q^{der}, similarly for G_q^{ad} and G_q^{sc}) and are in fact forms of G_0^{der}, G_0^{ad} and G_0^{sc} respectively (to see this note that $G_q^{\mathrm{der}} \times_F \Omega \cong (G_q \times_F \Omega)^{\mathrm{der}}$ and $Z(G_q) \times_F \Omega \cong Z(G_q \times_F \Omega)$). Using Proposition 13.1(1) of [Kottwitz 2014] and the fact that G_q^{ad} has trivial center, we have a canonical bijection

$$\kappa_{G_q} : H^1(\mathrm{Aut}(\Omega/F), G_q^{\mathrm{ad}}(\Omega)) \to (X_*(T_q^{\mathrm{ad}})/X_*(T_q^{\mathrm{sc}}))_{\mathrm{Aut}(\Omega/F)}.$$

Let $E_i(F, G_q)_m$ denote the F-isomorphism classes of inner forms of G_q that split over an at most m-ramified extension of F. Let (G'_q, B'_q, T'_q) correspond to the cocycle $q' \circ \mathfrak{Q}_m(s_{G_q})$ and let $E_i(F', G'_q)_m$ be the corresponding object over F'.

Lemma 5.1. *The congruence data D_m induces an isomorphism*

$$\mathfrak{I}_m : (X_*(T_q^{\mathrm{ad}})/X_*(T_q^{\mathrm{sc}}))_{\mathrm{Aut}(\Omega/F)} \xrightarrow{\sim} (X_*(T'^{\mathrm{ad}}_q)/X_*(T'^{\mathrm{sc}}_q))_{\mathrm{Aut}(\Omega'/F')}.$$

In particular, D_m induces a bijection $E_i(F, G_q)_m \to E_i(F', G'_q)_m$, $[G] \to [G']$ where $s_{G'} = \kappa_{G'_q}^{-1} \circ \mathfrak{I}_m \circ \kappa_{G_q}(s_G)$.

Proof. Note that $X_*(T_q) \cong X_*(T_0) \cong X_*(T'_q)$ as $\mathbb{Z}$-modules and the Galois action on $X_*(T_q)$ is determined by the cocycle s_{G_q} (and similarly for $X_*(T_q^{\mathrm{ad}})$, $X_*(T_q^{\mathrm{sc}})$). Now the lemma is obvious by Lemma 3.3. □

To proceed, we need to prove a version of Lemma 5.1 at the level of cocycles. To do this, we will use some results from Section 2 of [DeBacker and Reeder 2009].

Steinberg's vanishing theorem. Let G be a connected, reductive F-group. Steinberg's vanishing theorem asserts that

Theorem 5.2 [Steinberg 1965, Theorem 56]. $H^1(\mathrm{Gal}(F_s/F_{\mathrm{un}}), G(F_s)) = 1$.

As a corollary of this theorem, we obtain that the natural surjection from $\mathrm{Gal}(F_s/F) \to \mathrm{Gal}(F_{\mathrm{un}}/F)$ induces an isomorphism

$$H^1(\mathrm{Gal}(F_{\mathrm{un}}/F), G(F_{\mathrm{un}})) \cong H^1(\mathrm{Gal}(F_s/F), G(F_s)).$$

5.A. ***Congruence data for inner forms: a comparison of cocycles.*** Let A_q be a maximal F-split torus in G_q and let S_q be a maximal F_{un}-split F-torus in G_q that contains A_q. Let $T_q = Z_{G_q}(S_q)$. Then T_q is a maximal torus in $G_{q,F_{\mathrm{un}}}$ with maximal F_{un}-split torus S_q. Let C_q be an σ-stable alcove in $\mathcal{A}(S_q, F_{\mathrm{un}})$.

Let P_{C_q} be the Iwahori subgroup of $G_q^{\mathrm{ad}}(F_{\mathrm{un}})$ attached to C_q. Let $\Omega_{C_q}^{\mathrm{ad}} \subset \tilde{W}^{\mathrm{ad}} := X_*(T_q^{\mathrm{ad}})_{I_F} \rtimes W$ consist of elements which preserve the alcove C_q. Here I_F is the inertia subgroup of F and $W = W(G_{q,F_{\mathrm{un}}}, S_{q,F_{\mathrm{un}}})$. Then

$$\Omega_{C_q}^{\mathrm{ad}} \cong (X_*(T_q^{\mathrm{ad}})/X_*(T_q^{\mathrm{sc}}))_{I_F} \tag{5-1}$$

by Lemma 15 of [Haines and Rapoport 2008]. Let $P^*_{C_q}$ be the normalizer in G_q^{ad} of P_{C_q}. Let $N^{\mathrm{ad}}_{C_q} = N_{G_q^{\mathrm{ad}}}(S_q^{\mathrm{ad}})(F_{\mathrm{un}}) \cap P^*_{C_q}$. Then $\Omega^{\mathrm{ad}}_{C_q}$ is the image of $N^{\mathrm{ad}}_{C_q}$ in $\tilde{W}^{\mathrm{ad}}$ and $\Omega^{\mathrm{ad}}_{C_q} \cong P^*_{C_q}/P_{C_q}$.

The following lemma is proved in Sections 2.3 and 2.4 of [DeBacker and Reeder 2009]. Although the authors assume that $G_{q,F_{\mathrm{un}}}$ is split in the beginning of Section 2.3 of [DeBacker and Reeder 2009], this assumption is not needed in their proof of the following lemma. They use that when $G_{q,F_{\mathrm{un}}}$ is split, $\Omega^{\mathrm{ad}}_{C_q} \cong X_*(T_q^{\mathrm{ad}})/X_*(T_q^{\mathrm{sc}})$ in Corollary 2.4.2 and Corollary 2.4.3; one should instead use (5-1) when $G_{q,F_{\mathrm{un}}}$ is not necessarily split.

Lemma 5.3 [DeBacker and Reeder 2009, Corollary 2.4.3]. *We have isomorphisms*

$$H^1(\mathrm{Gal}(F_{\mathrm{un}}/F), \Omega^{\mathrm{ad}}_{C_q}) \cong H^1(\mathrm{Gal}(F_{\mathrm{un}}/F), N^{\mathrm{ad}}_{C_q}) \cong H^1(\mathrm{Gal}(F_{\mathrm{un}}/F), G_q^{\mathrm{ad}}(F_{\mathrm{un}})).$$

Let c be a cocycle in $Z^1(\mathrm{Gal}(F_{\mathrm{un}}/F), \Omega^{\mathrm{ad}}_{C_q})$. By Lemma 2.1.2 of [DeBacker and Reeder 2009], since $\Omega^{\mathrm{ad}}_{C_q}$ is finite, we have

$$Z^1(\mathrm{Gal}(F_{\mathrm{un}}/F), \Omega^{\mathrm{ad}}_{C_q}) = \Omega^{\mathrm{ad}}_{C_q}.$$

Let G be the inner form of G_q determined by c. Let $c(\sigma) = w_\sigma$. Write $w_\sigma = (\lambda, w)$ with $\lambda \in X_*(T^{\mathrm{ad}})_{I_F}$ and $w \in W$. Let $K \subset F_s$ denote the finite at most m-ramified extension of F_{un} over which $G_{q,F_{\mathrm{un}}}$ splits. Let $t = Nm(\tilde{\lambda}(\pi_K))$ where $Nm : T_q^{\mathrm{ad}}(K) \to T_q^{\mathrm{ad}}(F_{\mathrm{un}})$ and $\tilde{\lambda} \to \lambda$ under the usual surjection $X_*(T_q^{\mathrm{ad}}) \to X_*(T_q^{\mathrm{ad}})_I$. Let $\tilde{w} \in N_{G_q}(S_q)(F_{\mathrm{un}})$ be the representative of w chosen using the Chevalley–Steinberg system we fixed in Section 4.A.1.

Let $m_\sigma = t\tilde{w}$. Since w_σ stabilizes C_q, it follows that $m_\sigma P_{C_q} m_\sigma^{-1} = P_{C_q}$. Hence $m_\sigma \in P^*_{C_q}$. Therefore $\tilde{c}(\sigma) = m_\sigma \in Z^1(\mathrm{Gal}(F_{\mathrm{un}}/F), N^{\mathrm{ad}}_{C_q})$. Denoting

$$G(F_{\mathrm{un}}) \to G_q(F_{\mathrm{un}}), \quad g_* \to g,$$

the new action of σ on an element $g_* \in G(F_{\mathrm{un}})$, which we denote by σ_*, is given by

$$\sigma_* \cdot g_* = (\tilde{c}(\sigma)(\sigma \cdot g))_*$$

(Here $\sigma \cdot g$ denotes the action of σ on $g \in G_q(F_{\mathrm{un}})$). Note that $c(\sigma) \in G_q^{\mathrm{ad}}(F_{\mathrm{un}}) = \mathrm{Inn}(G_q)(F_{\mathrm{un}})$. The maximal F_{un}-split torus S_q of G_q gives a maximal F_{un}-split, F_{un}-torus S in G. Let $X_*(S) \to X_*(S_q)$, $\tau_* \to \tau$. For $\tau_* \in X_*(S)$, $\sigma_* \cdot \tau_* = (w_\sigma(\sigma \cdot \tau))_*$. Since S_q is defined over F, $\sigma \cdot \tau \in X_*(S_q)$. Since $w_\sigma \in \Omega^{\mathrm{ad}}_{C_q}$, we see that $X_*(S)$ is stable under the action of σ, and hence S is defined over F.

Lemma 5.4. *Let A be the F-split torus of G determined by the $\mathbb{Z}$-module $X_*(S)^{\sigma_*}$. Then A is a maximal F-split torus in G.*

Proof. Consider the reduced apartment $\mathcal{A}(S_q, \widehat{F_{\mathrm{un}}})$. We view this as an apartment in the reduced building of $G(\widehat{F_{\mathrm{un}}})$ and denote it as $\mathcal{A}(S, \widehat{F_{\mathrm{un}}})$. The action σ_* on $x_* \in \mathcal{A}(S, \widehat{F_{\mathrm{un}}})$ given by $\sigma_* \cdot x_* = (w_\sigma(\sigma \cdot x))_*$. Let C_* denote the alcove in $\mathcal{A}(S, \widehat{F_{\mathrm{un}}})$ corresponding to C_q. Then $\sigma_* \cdot C_* = (w_\sigma(\sigma \cdot C_q))_*$. Since $\sigma \cdot C_q = C_q$ and since $w_\sigma \in \Omega^{\mathrm{ad}}_{C_q}$, we see that C_* is a σ_*-stable alcove in $\mathcal{A}(S, \widehat{F_{\mathrm{un}}})$. In particular, $\mathcal{A}(S, \widehat{F_{\mathrm{un}}})$ is σ_*-stable. By Proposition 5.1.14 of [Bruhat and Tits 1984], $C_*^{\sigma_*}$ is an alcove in the affine space $\mathcal{A}(A, F)$. Since $\mathcal{A}(A, F)$ contains a facet of maximal possible dimension, we see that A is maximal F-split in G. □

Let (G'_q, T'_q, B'_q, S'_q) correspond to (G_q, T_q, B_q, S_q) via congruence data D_m as in Section 4. By Lemma 3.4, we have

$$\Omega^{\mathrm{ad}}_{C_q} \cong \Omega^{\mathrm{ad}}_{C'_q}.$$

Let $w_{\sigma'} \in \Omega^{\mathrm{ad}}_{C'_q}$ be the image of w_σ under this isomorphism. This isomorphism gives rise to a bijection of pointed sets

$$\begin{aligned} \mathfrak{I}_m : Z^1(\mathrm{Gal}(F_{\mathrm{un}}/F), \Omega^{\mathrm{ad}}_{C_q}) &\to Z^1(\mathrm{Gal}(F'_{\mathrm{un}}/F'), \Omega^{\mathrm{ad}}_{C'_q}), \\ c &\to c' \end{aligned} \tag{5-2}$$

where $c'(\sigma') = w_{\sigma'}$. Let $m_{\sigma'} = t'\tilde{w}'$ where $w_{\sigma'} = (\lambda', w') \in X_*(T^{\mathrm{ad}}_q)_{I_{F'}} \rtimes W'$. Here $t' = Nm(\tilde{\lambda}'(\pi'_{K'}))$ where $Nm : T^{\mathrm{ad}'}_q(K') \to T^{\mathrm{ad}'}_q(F'_{\mathrm{un}})$ and $\tilde{\lambda}' \to \lambda'$ under the usual surjection $X_*(T^{\mathrm{ad}'}_q) \to X_*(T^{\mathrm{ad}'}_q)_{I_{F'}}$, and $\tilde{\lambda} \to \tilde{\lambda}'$ under the isomorphism $X_*(T^{\mathrm{ad}}_q) \cong X_*(T^{\mathrm{ad}'}_q)$. Also $\tilde{w}'$ is the representative of w chosen using the Chevalley–Steinberg system fixed in Section 4.A.1. Let $\tilde{c}' \in Z^1(\mathrm{Gal}(F'_{\mathrm{un}}/F'), N^{\mathrm{ad}'}_{C'_q})$ be the cocycle with $\tilde{c}'(\sigma') = m_{\sigma'}$.

Let G' be the inner form of G'_q determined by c' (or $\tilde{c}'$). Let S' be the maximal F'_{un}-split, F_{un}-torus of G' corresponding to S'_q but with the action of σ' given by the cocycle $\tilde{c}'$. More precisely, for $g'_* \in G'(F'_{\mathrm{un}})$,

$$\sigma'_* \cdot g'_* = (\tilde{c}'(\sigma') \cdot (\sigma' \cdot g'))_*$$

where $\sigma' = \mathrm{Del}_m(\sigma)$ as before, and $\sigma' \cdot g'$ denotes the action of σ' on $G'_q(F'_{\mathrm{un}})$.

As in Lemma 5.4, we see that S' is an F'-torus that is maximal F'_{un}-split and whose split component A' is a maximal F'-split torus in G'.

Corollary 5.5. *With $G \to G'$ as above, the F-rank of G is equal to the F'-rank of G'.*

Proof. This is because $\mathrm{rank}(S) = \mathrm{rank}(S')$ and the isomorphism $X_*(S) \to X_*(S')$ is σ_*-equivariant. Hence $\mathrm{rank}(A) = \mathrm{rank}(A')$ by Lemma 5.4. □

6. Congruences of parahoric group schemes: étale descent

The following lemma is easy.

Lemma 6.1. *The σ-equivariant isomorphism $\tilde{A}_m : \mathcal{A}(S_q, \widehat{F_{\mathrm{un}}}) \to \mathcal{A}(S'_q, \widehat{F'_{\mathrm{un}}})$ induces a σ_*-equivariant isomorphism $\tilde{A}_{m,*} : \mathcal{A}(S, \widehat{F_{\mathrm{un}}}) \to \mathcal{A}(S', \widehat{F'_{\mathrm{un}}})$.*

Now let $\tilde{\mathcal{F}}_*$ be σ_*-invariant facet in $\mathcal{A}(S, \widehat{F_{\mathrm{un}}})$ and let $\tilde{\mathcal{F}}'_* = \tilde{A}_{m,*}(\tilde{\mathcal{F}}_*)$. Let $\mathcal{F}_* = \tilde{\mathcal{F}}^{\sigma_*}_*$ and $\mathcal{F}'_* = \tilde{\mathcal{F}}'^{\sigma'_*}$.

Proposition 6.2. *Let $m \geq 1$, F, F' nonarchimedean local fields with $e_F, e_{F'} \geq 2m$. Let l as in Theorem 4.5, let D_l be the congruence data of level l, and let (G'_q, T'_q, B'_q, S'_q) correspond to (G_q, T_q, B_q, S_q) via D_l. Let $\tilde{p}_m : \mathcal{P}_{\tilde{\mathcal{F}}} \times_{\mathfrak{O}_{\widehat{F_{\mathrm{un}}}}} \mathfrak{O}_{\widehat{F_{\mathrm{un}}}}/\mathfrak{p}^m_{\widehat{F_{\mathrm{un}}}} \to \mathcal{P}_{\tilde{\mathcal{F}}'} \times_{\mathfrak{O}_{\widehat{F'_{\mathrm{un}}}}} \mathfrak{O}_{\widehat{F'_{\mathrm{un}}}}/\mathfrak{p}^m_{\widehat{F'_{\mathrm{un}}}}$ denote the σ-equivariant isomorphism of Theorem 4.5 and Proposition 4.10. Let $c \to c'$ via $\mathfrak{I}_m$ (see (5-2)). The isomorphism $\tilde{p}_m$ induces a σ_*-equivariant isomorphism $\tilde{p}_{m,*} : \mathcal{P}_{\tilde{\mathcal{F}}_*} \times_{\mathfrak{O}_{\widehat{F_{\mathrm{un}}}}} \mathfrak{O}_{\widehat{F_{\mathrm{un}}}}/\mathfrak{p}^m_{\widehat{F_{\mathrm{un}}}} \to \mathcal{P}_{\tilde{\mathcal{F}}'_*} \times_{\mathfrak{O}_{\widehat{F'_{\mathrm{un}}}}} \mathfrak{O}_{\widehat{F'_{\mathrm{un}}}}/\mathfrak{p}^m_{\widehat{F'_{\mathrm{un}}}}$.*

Proof. We begin by understanding the action of σ_* on an element of $P_{\tilde{\mathcal{F}}_*}$ more explicitly. Recall that

$$P_{\tilde{\mathcal{F}}} = \langle \mathcal{U}^+_{\tilde{\mathcal{F}}}(\mathfrak{O}_{\widehat{F_{\mathrm{un}}}}), \mathcal{T}(\mathfrak{O}_{\widehat{F_{\mathrm{un}}}}), \mathcal{U}^-_{\tilde{\mathcal{F}}}(\mathfrak{O}_{\widehat{F_{\mathrm{un}}}})\rangle$$

Let $g \in P_{\tilde{\mathcal{F}}}$. Then $\sigma_* \cdot g_* = (m_\sigma(\sigma \cdot g)m_\sigma^{-1})_*$. Let $b_0 \in \Phi^{\mathrm{red}}(G_q, S_q)$ such that $2b_0$ is not a root. Let $y \in U_{b_0,\tilde{\mathcal{F}}}$. Fix $\beta_0|_{S_q} = b_0$, fix the pinning (L_{β_0}, x_{b_0}) and write $y = x_{b_0}(u_0)$ for $u_0 \in L_{\beta_0}$ (As explained in Section 2.C.2, $L_{b_0} \cong L_{\beta_0} \hookrightarrow K$). Let $\tilde{\sigma}$ denote a lift of σ to Γ_F and let $\beta = \tilde{\sigma} \cdot \beta_0$, $b = \sigma \cdot b_0$. Then we obtain a pinning (L_β, x_b) from the pinning (L_{β_0}, x_{b_0}) via $\tilde{\sigma}$ and we have $\sigma \cdot x_{b_0}(u_0) = x_b(\tilde{\sigma} \cdot u_0)$; this follows using properties of Chevalley–Steinberg system recalled in Section 2.C.1 (a), (b). Let $u = \tilde{\sigma} \cdot u_0$. Then $u \in L_\beta$. We need to compute $\tilde{w} x_b(u)\tilde{w}^{-1}$. We will first compute $\tilde{s}_a x_b(u)\tilde{s}_a^{-1}$ for $a \in \Delta$. Note that

$$\tilde{s}_a = \prod_{\alpha \in \tilde{\Delta}_a} \tilde{s}_\alpha \tag{6-1}$$

and that $L_{s_a \cdot b} = L_b$. Now for $\alpha_1, \beta_1 \in \Phi(G_q, T_q)$, we have $\tilde{s}_{\alpha_1} x_{\beta_1}(z)\tilde{s}_{\alpha_1}^{-1} = x_{s_{\alpha_1}(\beta_1)}(d_{\alpha_1,\beta_1} z)$ for all $z \in K$, with $d_{\alpha_1,\beta_1} = \pm 1$. Using the properties of Chevalley–Steinberg system recalled in Section 2.C.1 (a), (b), we have

$$d_{\alpha_1,\beta_1} = d_{\gamma(\alpha_1),\gamma(\beta_1)} \quad \forall \gamma \in \mathrm{Gal}(K/\widehat{F_{\mathrm{un}}}). \tag{6-2}$$

With β as above, note that $\beta|_{S_q} = b$. Let

$$d_{a,b} := \prod_{\alpha \in \tilde{\Delta}_a} d_{\alpha,\beta}$$

This notation is justified since (6-2) implies that the definition of $d_{a,b}$ does not depend on the choice of β. Using the definition of x_b in (2-2), a simple calculation yields that $\tilde{s}_a x_b(u)\tilde{s}_a^{-1} = x_{s_a(b)}(d_{a,b}u)$. Since we chose our Chevalley–Steinberg systems compatibly (see Section 4.A.1), we evidently have $d_{a,b} = d_{a',b'}$ for all $a \in \Delta$, $b \in \Phi(G_q, S_q)$. Iterating this process, we see that $\tilde{w} x_b(u)\tilde{w}^{-1} = x_{w \cdot b}(d_{w,b}u)$ where $d_{w,b} = \pm 1$ and $d_{w,b} = d_{w',b'}$.

Suppose $b_0 \in \Phi(G_q, S_q)$ such that $2b_0$ is a root. Let $\beta_0, \bar{\beta}_0|_{S_q} = b_0$. Fix the pinning $(L_{\beta_0}, L_{\beta_0+\bar{\beta}_0}, x_{b_0})$ and write $y = x_{b_0}(u_0, v_0)$, with $u_0, v_0 \in L_{\beta_0}$ (Recall that $L_{b_0} \cong L_{\beta_0} \subset K$). Let $\beta = \tilde{\sigma} \cdot \beta_0$, $\bar{\beta} = \tilde{\sigma} \cdot \bar{\beta}_0$ and $b = \sigma \cdot b_0$. We then obtain a pinning $(L_\beta, L_{\beta+\bar{\beta}}, x_b)$ via $\tilde{\sigma}$ and $\sigma \cdot x_{b_0}(u_0, v_0) = x_{\sigma \cdot b_0}(\tilde{\sigma} \cdot u_0, \tilde{\sigma} \cdot v_0)$ where $\tilde{\sigma}$ as before. Let $u = \tilde{\sigma} \cdot u_0$, $v = \tilde{\sigma} \cdot v_0$. Then $u, v \in L_\beta$. We need to compute $\tilde{s}_a x_b(u, v)\tilde{s}_a^{-1}$ where s_a is as in (6-1). Let

$$d_{a,b} := \prod_{\alpha \in \tilde{\Delta}_a} d_{\alpha,\beta}, \quad d_{a,2b} := \prod_{\alpha \in \tilde{\Delta}_a} d_{\alpha,\beta+\bar{\beta}}$$

Again, the definitions of $d_{a,b}$ and $d_{a,2b}$ do not depend on the choice of β by (6-2).

Then a simple calculation yields

$$\tilde{s}_a x_b(u, v)\tilde{s}_a^{-1} = x_{s_a(b)}(d_{a,b}u, d_{a,2b}v).$$

Proceeding as in the previous case, we have $\tilde{w} x_b(u, v)\tilde{w}^{-1} = x_{w \cdot b}(d_{w,b}u, d_{w,2b}v))$ where $d_{w,b}, d_{w,2b} = \pm 1$ and $d_{w,b} = d_{w',b'}$ and $d_{w,2b} = d_{w',2b'}$.

Recall that $t = Nm(\tilde{\lambda}(\pi_K)) \in T_q^{\mathrm{ad}}(F_{\mathrm{un}})$. Then for each $\gamma \in \mathrm{Gal}(K/F_{\mathrm{un}})$, $\gamma \cdot t = t$. Let $c \in \Phi^{\mathrm{red}}(G_q, S_q)$ with $2c$ not a root. Let $\chi \in \Phi(G_q.T_q)$ with $\chi|_{S_q} = c$. Note that χ factors through T_q^{ad}. Fixing the pinning (L_χ, x_c) we have that $\chi : T \to G_m$ is defined over L_χ and $\chi(t) \in L_\chi^\times$. A simple calculation yields $tx_c(u)t^{-1} = x_c(\chi(t)u)$ for each $u \in L_\chi$. If c, $2c$ are roots, then with $\chi, \bar{\chi}$ such that $\chi|_{S_q} = \bar{\chi}|_{S_q} = c$ and fixing the pinning $(L_\chi, L_{\chi+\bar{\chi}}, x_c)$, it follows that $tx_c(u, v)t^{-1} = x_c(\chi(t)u, (\chi+\bar{\chi})(t)v)$. Hence, if $2b_0$ is not a root then

$$\sigma_* \cdot (x_{b_0}(u_0))_* = (x_{w \cdot b}(d_{w,b}\chi(t)u))_*$$

where $\chi|_{S_q} = w \cdot b$. If $2b_0$ is a root, then

$$\sigma_* \cdot (x_{b_0}(u_0, v_0))_* = (x_{w \cdot b}(d_{w,b}\chi(t)u, d_{w,2b}(\chi+\bar{\chi})(t)v))_*$$

where $\chi, \chi' \in \Phi(G_q, T_q)$ are such that $\chi, \bar{\chi}|_{S_q} = w \cdot b$. It is easy to calculate $\sigma_* \cdot (x_{2b_0}(0, v_0))_*$ using the observations above. For $x \in \mathcal{T}_q(\mathfrak{O}_{\widehat{F_{\mathrm{un}}}})$,

$$\sigma_* \cdot x_* = (w(\sigma \cdot x)w^{-1})_*.$$

Combining these observations with the fact that $\tilde{p}_m$ is σ-equivariant (see Proposition 4.10), it follows that the map $\tilde{p}_{m,*}$ has the property that $\tilde{p}_{m,*} \circ \sigma_* = \sigma'_* \circ \tilde{p}_{m,*}$ (in this verification, we choose $\tilde{\sigma}'$ to correspond to $\tilde{\sigma}$ via Del_m). □

Corollary 6.3. *The isomorphism $\tilde{p}_{m,*}$ induces an isomorphism of group schemes*

$$p_{m,*} : \mathcal{P}_{\mathcal{F}_*} \times_{\mathfrak{O}_F} \mathfrak{O}_F/\mathfrak{p}_F^m \to \mathcal{P}_{\mathcal{F}'_*} \times_{\mathfrak{O}_{F'}} \mathfrak{O}_{F'}/\mathfrak{p}_{F'}^m \times_{\psi_m^{-1}} \mathfrak{O}_F/\mathfrak{p}_F^m.$$

In particular $\mathcal{P}_{\mathcal{F}_*}(\mathfrak{O}_F/\mathfrak{p}_F^m)$ and $\mathcal{P}_{\mathcal{F}'_*}(\mathfrak{O}_{F'}/\mathfrak{p}_{F'}^m)$ are isomorphic as groups.

Proof. This follows from Proposition 6.2 and étale descent [Bosch et al. 1990, Example B, Section 6.2]. □

Acknowledgments

I would like to express my gratitude to J. K. Yu for introducing me to questions related to this article and for the insightful discussions during my graduate school years. I thank the referees for the helpful comments and suggestions, particularly regarding removing the assumption on residue characteristic that was made in an earlier version of this article. I thank Dipendra Prasad and K. V. Shuddhodan for the interesting discussions and for their encouragement.

References

[Aubert et al. 2016] A.-M. Aubert, P. Baum, R. Plymen, and M. Solleveld, "The local Langlands correspondence for inner forms of SL_n", *Res. Math. Sci.* **3** (2016), art. id. 32. MR Zbl

[Badulescu 2002] A. I. Badulescu, "Correspondance de Jacquet–Langlands pour les corps locaux de caractéristique non nulle", *Ann. Sci. École Norm. Sup.* (4) **35**:5 (2002), 695–747. MR Zbl

[Bosch et al. 1990] S. Bosch, W. Lütkebohmert, and M. Raynaud, *Néron models*, Ergebnisse der Mathematik (3) **21**, Springer, 1990. MR Zbl

[Bruhat and Tits 1972] F. Bruhat and J. Tits, "Groupes réductifs sur un corps local", *Inst. Hautes Études Sci. Publ. Math.* **41** (1972), 5–251. MR Zbl

[Bruhat and Tits 1984] F. Bruhat and J. Tits, "Groupes réductifs sur un corps local, II: Schémas en groupes, existence d'une donnée radicielle valuée", *Inst. Hautes Études Sci. Publ. Math.* **60** (1984), 197–376. MR Zbl

[Chai and Yu 2001] C.-L. Chai and J.-K. Yu, "Congruences of Néron models for tori and the Artin conductor", *Ann. of Math.* (2) **154**:2 (2001), 347–382. MR Zbl

[Conrad 2011] B. Conrad, "Reductive group schemes", preprint, SGA3 Summer School, 2011, Available at https://tinyurl.com/conreduct.

[DeBacker and Reeder 2009] S. DeBacker and M. Reeder, "Depth-zero supercuspidal L-packets and their stability", *Ann. of Math.* (2) **169**:3 (2009), 795–901. MR Zbl

[Deligne 1984] P. Deligne, "Les corps locaux de caractéristique p, limites de corps locaux de caractéristique 0", pp. 119–157 in *Representations of reductive groups over a local field*, Hermann, Paris, 1984. MR Zbl

[Ganapathy 2015] R. Ganapathy, "The local Langlands correspondence for GSp_4 over local function fields", *Amer. J. Math.* **137**:6 (2015), 1441–1534. MR Zbl

[Ganapathy and Varma 2017] R. Ganapathy and S. Varma, "On the local Langlands correspondence for split classical groups over local function fields", *J. Inst. Math. Jussieu* **16**:5 (2017), 987–1074. MR Zbl

[Haines and Rapoport 2008] T. Haines and M. Rapoport, "Appendix: On parahoric subgroups", 2008. Appendix to G. Pappas and M. Rapoport, "Twisted loop groups and their affine flag varieties", *Adv. Math.* **219**:1 (2008), 118–198. MR Zbl

[Howe 1985] R. Howe, *Harish-Chandra homomorphisms for p-adic groups*, CBMS Regional Conf. Series in Math. **59**, Amer. Math. Soc., Providence, RI, 1985. MR Zbl

[Kazhdan 1986] D. Kazhdan, "Representations of groups over close local fields", *J. Analyse Math.* **47** (1986), 175–179. MR Zbl

[Kottwitz 2014] R. Kottwitz, "B(G) for all local and global fields", preprint, 2014. arXiv

[Landvogt 1996] E. Landvogt, *A compactification of the Bruhat–Tits building*, Lecture Notes in Math. **1619**, Springer, 1996. MR Zbl

[Lemaire 2001] B. Lemaire, "Représentations génériques de GL_N et corps locaux proches", *J. Algebra* **236**:2 (2001), 549–574. MR Zbl

[Serre 1979] J.-P. Serre, *Local fields*, Graduate Texts in Math. **67**, Springer, 1979. MR Zbl

[Steinberg 1965] R. Steinberg, "Regular elements of semisimple algebraic groups", *Inst. Hautes Études Sci. Publ. Math.* **25** (1965), 49–80. MR Zbl

Communicated by Michael Rapoport
Received 2018-11-06 Revised 2019-04-16 Accepted 2019-05-25

radhikag@iisc.ac.in *School of Mathematics, Tata Institute of Fundamental Research, Mumbai, India*

ALGEBRA AND NUMBER THEORY 13:6 (2019)
dx.doi.org/10.2140/ant.2019.13.1501

An improved bound for the lengths of matrix algebras

Yaroslav Shitov

Let S be a set of $n \times n$ matrices over a field $\mathbb{F}$. We show that the $\mathbb{F}$-linear span of the words in S of length at most

$$2n \log_2 n + 4n$$

is the full $\mathbb{F}$-algebra generated by S. This improves on the $\frac{n^2}{3} + \frac{2}{3}$ bound by Paz (1984) and an $O(n^{3/2})$ bound of Pappacena (1997).

Let S be a subset of a finite-dimensional associative algebra $\mathcal{A}$ over a field $\mathbb{F}$. An element $a \in \mathcal{A}$ is said to be a *word* of length k in S if there are $a_1, \dots, a_k \in S$ such that $a = a_1 \cdots a_k$. We denote the set of all such words by S^k, and we write $\mathbb{F}S^k$ for the $\mathbb{F}$-linear span of S^k. Similarly, $\mathbb{F}S^{\leqslant k}$ will stand for the $\mathbb{F}$-linear span of all the words in S that have length at most k.

Definition 1. The length $\ell(S)$ is the smallest integer k for which $\mathbb{F}S^{\leqslant k}$ is the full subalgebra generated by S. We also define $\ell(\mathcal{A})$ as the maximum value of $\ell(S)$, where S runs over all subsets of $\mathcal{A}$ that generate $\mathcal{A}$ as an $\mathbb{F}$-algebra.

In our paper, we study the length of $\mathrm{Mat}_n(\mathbb{F})$, the set of $n \times n$ matrices viewed as an algebra over $\mathbb{F}$. A. Paz [1984] proved that $\ell(S) \leqslant \frac{n^2}{3} + \frac{2}{3}$ for all $S \subset \mathrm{Mat}_n(\mathbb{F})$ and proposed the following appealing conjecture.

Conjecture 2. *For all* $S \subset \mathrm{Mat}_n(\mathbb{F})$, *one has* $\ell(S) \leqslant 2n - 2$.

As shown by T. Laffey [1986, page 131], the upper bound in Conjecture 2 should be sharp. This conjecture is known to hold if the size of matrices is at most four [Paz 1984] or if $\mathbb{F}S$ contains a nonderogatory matrix [Guterman et al. 2018]. However, the best known general upper bounds on the lengths of matrix subsets are quite far from the one prescribed by Conjecture 2. It was only in 1997 when a subquadratic estimate was obtained: C. Pappacena proved an $O(n^{3/2})$ upper bound on the length of $\mathrm{Mat}_n(\mathbb{F})$, but no further improvements have been made since then [Guterman et al. 2018; Lambrou and Longstaff 2009; Longstaff et al. 2006]. The main result of this paper is a much stronger $O(n \log n)$ upper bound on the length of $\mathrm{Mat}_n(\mathbb{F})$.

Theorem 3. *For all* $S \subset \mathrm{Mat}_n(\mathbb{F})$, *we have* $\ell(S) \leqslant 2n \log_2 n + 4n - 4$.

MSC2010: primary 15A03; secondary 15A30.
Keywords: matrix theory, finite-dimensional algebras, generating sets.

As an additional motivation of our study, we note that the best known upper bounds on a complete set of unitary invariants for $n \times n$ matrices [Laffey 1986] and on the PI degree of semiprime affine algebras of Gelfand–Kirillov dimension one [Pappacena et al. 2003] come from the estimates of $\ell(\mathrm{Mat}_n(\mathbb{F}))$, so the current work also improves our understanding of those invariants.

1. Warm-up

In this section, we explain the idea behind our main construction and illustrate how it works in a simpler setting. We get a small improvement on one of the results of Pappacena [1997], which allows us to prove the $n = 5$ case of Conjecture 2.

We say that a set $S \subset \mathrm{Mat}_n(\mathbb{F})$ is *irreducible* if it generates $\mathrm{Mat}_n(\mathbb{F})$ as the $\mathbb{F}$-algebra. If a set S is not irreducible, and if $\mathbb{F}$ is algebraically closed, then there exist $p \in \{1, \dots, n-1\}$ and $Q \in \mathrm{GL}_n(\mathbb{F})$ such that, for any $A \in S$, we have

$$Q^{-1} A Q = \begin{pmatrix} A_{11} & A_{12} \\ O & A_{22} \end{pmatrix} \tag{1-1}$$

where A_{11} is a $p \times p$ matrix (and O is the zero matrix of appropriate dimensions). This is *Burnside's theorem*; see [Radjavi and Rosenthal 2000, Theorem 1.5.1].

Lemma 4 [Markova 2005, Corollary 3]. *Let $\mathcal{A}$ be a matrix algebra whose elements are of the form (1-1), and let $\mathcal{A}_1, \mathcal{A}_2$ be the sets of all A_{11}, A_{22} blocks of matrices in $\mathcal{A}$, respectively. Then $\ell(\mathcal{A}) \leqslant \ell(\mathcal{A}_1) + \ell(\mathcal{A}_2) + 1$.*

We will say that a matrix $Z \in \mathrm{Mat}_n(F)$ is *square-zero* if $Z^2 = 0$. The main idea of the proof of Theorem 3 is to control the product $\lambda\rho(\lambda)$, where $\rho(\lambda)$ is the minimal rank of nonzero square-zero matrices that arise as linear combinations of words of length at most λ. We show in Section 2 below that we can reduce ρ to 1 whilst saving the property $\lambda\rho(\lambda) \in O(n \log n)$, and then we apply Pappacena's technique to deal with low rank matrices; see [Pappacena 1997, Theorem 4.1] and Corollary 7 below. More precisely, let $H \in \mathbb{F}S^{\leqslant \lambda}$ be a square-zero matrix; it can be written as

$$H = \begin{pmatrix} O & O & I_\rho \\ O & O & O \\ O & O & O \end{pmatrix}$$

with respect to some basis. If some matrix A with bottom-left block of small rank $r > 0$ comes as a linear combination of words of length l, then the matrix HAH is square-zero, has rank r, and comes as a linear combination of words of length at most $l + 2\lambda$. As we will see in Claims 13 and 14 below, we can always find an appropriate matrix A to reduce the rank of a square-zero matrix. The following lemma illustrates our approach to the proof of Claim 13.

Lemma 5. *Consider an irreducible set $S \subset \mathbb{F}^{n \times n}$ and a nonzero vector $v \in \mathbb{F}^n$. If $\mathbb{F}S^{\leqslant (n-2)} v \neq \mathbb{F}^n$, then $\mathbb{F}S$ contains a matrix with minimal polynomial of degree n.*

Proof. The sequence

$$\mathbb{F}v = \mathbb{F}S^0 v \subset \mathbb{F}S^{\leqslant 1} v \subset \cdots \subset \mathbb{F}S^{\leqslant k} v = \mathbb{F}^n$$

is strictly increasing [Pappacena 1997, Theorem 4.1], so the assumption of the lemma implies $k = n-1$ and $\dim \mathbb{F}S^{\leqslant t} v - \dim \mathbb{F}S^{\leqslant (t-1)} v = 1$ for all $t \in \{1, \dots, n-1\}$. Therefore, we can set $\mathcal{B}_0 = \{v\}$ and inductively complete $\mathcal{B}_{t-1}$ to a basis $\mathcal{B}_t$ of $\mathbb{F}S^{\leqslant t}$ by adding a single vector v_t. With respect to the basis $\{v, v_1, \dots, v_{n-1}\}$, every matrix in S has the form

$$A = \begin{pmatrix} * & \cdots & \cdots & * & * \\ a_{21} & * & \cdots & * & * \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & * & * \\ 0 & \cdots & 0 & a_{n,n-1} & * \end{pmatrix}$$

with $*$'s denoting the entries we need not specify. Since S is irreducible, all of the $(i+1, i)$ entries are nonzero at some matrix in S, so a generic element of $\mathbb{F}S$ has all of them nonzero — which means that its minimal polynomial has degree n. □

Theorem 6 [Guterman et al. 2018, Theorems 2.4 and 2.5]. *If an irreducible set $\mathbb{F}S \subset \mathrm{Mat}_n(\mathbb{F})$ contains a matrix with minimal polynomial of degree $n-1$ or n, then $\ell(S) \leqslant 2n-2$.*

Lemma 5 and Theorem 6 lead to a tiny improvement of the $r = 1$ case of Theorem 4.1(a) in [Pappacena 1997], which is nevertheless useful to study the case of small n.

Corollary 7. *Let $S \subset \mathrm{Mat}_n(\mathbb{F})$ be an irreducible set and $k \geqslant 2$. If $\mathbb{F}S^{\leqslant k}$ contains a rank-one matrix, then $\ell(S) \leqslant 2n+k-4$.*

Proof. If $\mathbb{F}S$ contains a matrix with minimal polynomial of degree n, then we are done by Theorem 6. Otherwise, we use Lemma 5 and get

$$\mathbb{F}S^{\leqslant (n-2)} A S^{\leqslant (n-2)} = \sum \mathrm{Mat}_n(\mathbb{F}) \cdot A \cdot \mathrm{Mat}_n(\mathbb{F}) = \mathrm{Mat}_n(\mathbb{F})$$

for any rank-one matrix A. □

We are almost ready to prove the $n = 5$ case of Conjecture 2.

Claim 8. *Assume that the minimal polynomial of every matrix in $\mathbb{F}S \subset \mathrm{Mat}_n(\mathbb{F})$ has degree at most 2. Then $\ell(S) \leqslant 2\log_2 n$.*

Proof. We denote by w a word in $S^{\ell(S)}$ that is not spanned by shorter words. For any $A, B \in S$, the matrices A^2 and $AB + BA = (A+B)^2 - A^2 - B^2$ belong to $\mathbb{F}S^{\leqslant 1}$, which implies that the letters of w are all different and their permutations do not break the property of w not to be spanned by shorter words. In particular, the products corresponding to the different $2^{\ell(S)}$ subsets of letters of w should be linearly independent, which implies $2^{\ell(S)} \leqslant \dim \mathrm{Mat}_n(\mathbb{F})$. □

Theorem 9. *If $S \subset \mathrm{Mat}_5(\mathbb{F})$, then $\ell(S) \leqslant 8$.*

Proof. Since a set of vectors is linearly dependent over $\mathbb{F}$ if it is linearly dependent over the algebraic closure of $\mathbb{F}$, it is sufficient to prove the statement assuming that $\mathbb{F}$ is algebraically closed [Guterman et al. 2018, page 239]. Moreover, Conjecture 2 is known to hold for $n \leqslant 4$ (see [Paz 1984]), so we can use Lemma 4 and assume without loss of generality that S is irreducible. According to Theorem 6 and Claim 8, we can restrict to the case when $\mathbb{F}S$ contains a matrix A with minimal polynomial of degree 3. A straightforward analysis of possible Jordan forms of A shows that the linear span of I, A, A^2 must contain a rank-one matrix, so it remains to apply Corollary 7. □

As said above, the case of $n \leqslant 4$ in Conjecture 2 was considered by Paz [1984], but the case of $n = 5$ remained open until now [Guterman et al. 2018]. Let us mention the works [Lambrou and Longstaff 2009; Longstaff et al. 2006], which cover the case $n \leqslant 6$ under the additional assumption of $\dim \mathbb{F}S \leqslant 2$.

2. The proof of Theorem 3

Let A be an $n \times n$ matrix over a field $\mathbb{F}$, which is assumed to be algebraically closed in this section. We recall that there exists $Q \in \mathrm{GL}_n(\mathbb{F})$ such that $Q^{-1}AQ$ has *rational normal form*, that is, we have $Q^{-1}AQ = \mathrm{diag}(C_{f_1}, \ldots, C_{f_k})$, where

$$C_f = \begin{pmatrix} 0 & 0 & \cdots & 0 & -c_0 \\ 1 & 0 & \cdots & 0 & -c_1 \\ 0 & 1 & \cdots & 0 & -c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -c_{m-1} \end{pmatrix}$$

is the companion matrix of a polynomial $f = t^m + c_{m-1}t^{m-1} + \cdots + c_0$, and the *invariant factors* $f_1, \cdots, f_k$ satisfy $f_1 | \cdots | f_k$.

Claim 10. *Let δ be the degree of the minimal polynomial of an $n \times n$ matrix A over $\mathbb{F}$. Then the $\mathbb{F}$-linear span of $I, A, \ldots, A^{\delta-1}$ contains either a nonzero projector of rank at most n/δ or a nonzero square-zero matrix of rank at most n/δ.*

Proof. We recall that the minimal polynomial φ of A occurs (one or more times) as an invariant factor of A. Let ψ be a polynomial that has degree $\delta - 1$, divides φ and is a multiple of any invariant factor different from φ. Then $\psi(A)$ has equal rank-one matrices in the places of the largest blocks of the rational normal form of A and zeros everywhere else. □

Claim 11. *For any irreducible set $S \subset \mathrm{Mat}_n(\mathbb{F})$, there exist nonzero λ, ρ such that $\lambda\rho \leqslant 2n$ and $\mathbb{F}S^{\leqslant\lambda}$ contains a square-zero matrix of rank ρ.*

Proof. We apply Claim 10 to any nonscalar matrix in S and find a nonzero matrix $P \in \mathbb{F}S^{\leqslant(\delta-1)}$ that has rank at most n/δ and satisfies either $P^2 = P$ or $P^2 = 0$. We are done if $P^2 = 0$; otherwise $H_B = (I - P)BP$ is a square-zero matrix for all B. We can have $H_B = 0$ only when the columns of BP are in the kernel of $I - P$, but this kernel being equal to $\mathrm{Im}\, P$ should then be invariant with respect to B, but since S is irreducible, this obstruction cannot happen for all $B \in S$. □

Claim 12. *Let* $A \in \mathbb{F}^{n\times n}$ *and* $r \in \mathbb{N}$. *Assume that the inequality* $\operatorname{rank}(PAQ) \leqslant r$ *holds, with any positive integers* p, q, *for all matrices* $P \in \mathbb{F}^{p\times n}$, $Q \in \mathbb{F}^{n\times q}$ *satisfying* $PQ = 0$. *Then* $\operatorname{rank}(A - \mu I) \leqslant 2r$ *for some* $\mu \in \mathbb{F}$.

Proof. Both the assumption and conclusion are independent of the substitution $A \to C^{-1}AC$, so we can assume that A has rational normal form. We denote the number of diagonal blocks by k and their sizes by $m_1, \ldots, m_k$. Since the ranks of the diagonal blocks cannot decrease by more than one upon adding a scalar matrix, and since the characteristic polynomials of these blocks have a common factor, we have $\min_\mu \operatorname{rank}(A - \mu I) = n - k$. We conclude the proof by constructing an identity square submatrix $A' = A[I|J]$ with $I \cap J = \varnothing$ and $|I| = |J| \geqslant 0.5(n-k)$, which would allow us to define P and Q as having the identity matrices at the $I \times I$ and $J \times J$ blocks and completed by an appropriate number of zero columns and rows, respectively. Namely, we pick a family of $\lfloor m_t/2 \rfloor$ nonconsecutive subdiagonal ones from a t-th diagonal block of A, and the union of all such families will be the diagonal of A'. □

Claim 13. *Let* $S \subset \mathbb{F}^{n\times n}$, $P \in \mathbb{F}^{p\times n}$, $Q \in \mathbb{F}^{n\times q}$. *Let* k *be the smallest integer such that* $PS^kQ \neq 0$. *Then, for any* $A_1, \ldots, A_k \in S$, *we have* $\operatorname{rank}(PA_1 \cdots A_k Q) \leqslant n/k$.

Proof. Let $V_0 = \operatorname{Im} Q$ and $V_t = \sum_{M \in S^{\leqslant t}} \operatorname{Im} MQ$. Let $\mathcal{B}_0, \ldots, \mathcal{B}_k \subset \mathbb{F}^n$ be vector families such that $\mathcal{B}_0 \cup \cdots \cup \mathcal{B}_t$ is a basis of V_t for $t = 0, \ldots, k$. Let $\mathcal{C} \subset \mathbb{F}^n$ be such that $\mathcal{B}_0 \cup \cdots \cup \mathcal{B}_k \cup \mathcal{C}$ is a basis of $\mathbb{F}^n$. Every matrix $A \in S$ has the form

$$\begin{pmatrix} & \mathcal{B}_0 & \mathcal{B}_1 & \cdots & \mathcal{B}_{k-1} & \mathcal{B}_k & \mathcal{C} \\ \mathcal{B}_0 & * & \cdots & \cdots & \cdots & * & * \\ \mathcal{B}_1 & A(1,0) & * & \cdots & \cdots & * & * \\ \mathcal{B}_2 & O & A(2,1) & * & \cdots & * & * \\ \vdots & \vdots & O & \ddots & * & \vdots & \vdots \\ \mathcal{B}_k & \vdots & \vdots & \ddots & A(k,k-1) & * & * \\ \mathcal{C} & O & O & \cdots & O & * & * \end{pmatrix},$$

where the $*$'s stand for entries that we need not specify, and the left column and top row of the matrix above indicate the basis vectors the respective blocks of rows and columns correspond to. We also have $P = (O|\cdots|O|P'|*)$, $Q = (\mathcal{Q}^\top|O|\cdots|O)^\top$ with some matrices P', $\mathcal{Q}$ at the $\mathcal{B}_k$ position of P and the $\mathcal{B}_0$ position of Q, respectively. For $A_1, \ldots, A_k \in S$, the matrix $PA_k \cdots A_1 Q$ equals $P'A_k(k, k-1) \cdots A_1(1,0)\mathcal{Q}$, so its rank is at most the smallest dimension of any of the matrices $A_k(k, k-1), \ldots, A_1(1,0)$, which is $\min_t |\mathcal{B}_t| \leqslant n/k$. □

Claim 14. *Let* $S \subset \operatorname{Mat}_n(\mathbb{F})$ *be an irreducible set and assume that* $\mathbb{F}S^{\leqslant\lambda}$ *contains a square-zero matrix* H *of rank* $\rho \geqslant 2$. *Then there exist* $\rho_1 \in [1, 0.5\rho]$ *and*

$$\lambda_1 \leqslant \frac{\lambda\rho}{\rho_1} + \frac{4n(\rho - \rho_1)}{\rho\rho_1}$$

such that $\mathbb{F}S^{\leqslant\lambda_1}$ *contains a square-zero matrix of rank equal to* ρ_1.

Proof. Let $P \in \mathbb{F}^{p\times\rho}$, $Q \in \mathbb{F}^{\rho\times q}$ be nonzero matrices satisfying $PQ = 0$. We choose a basis such that

$$H = \begin{pmatrix} O & O & I_\rho \\ O & O & O \\ O & O & O \end{pmatrix}$$

and define $P' = (O|O|P)$ and $Q' = (Q^\top|O|O)^\top$. Let k be the smallest integer for which there exist p, q and matrices P', Q' defined as above, and also $A_1, \dots, A_k \in S$ satisfying $P'A_1 \dots A_k Q' \neq 0$ (such an integer k exists because S is irreducible). We write $A = A_1 \cdots A_k$, and we denote by A' the bottom left block of A. Since $PA'Q \neq 0$, the matrix A' is nonscalar, that is, its minimal polynomial has degree $\delta > 1$.

Case 1. Assume $k \leqslant 4n/\rho$. By Claim 10, there is a polynomial ψ of degree at most $(\delta - 1)$ such that $\rho_1 := \operatorname{rank} \psi(A') \in [1, \rho/\delta]$; we see that $H_1 = \psi(HA)H$ is a square-zero matrix of rank ρ_1. It remains to note that H_1 is spanned by words of length at most

$$(\delta - 1)(\lambda + k) + \lambda \leqslant \lambda\delta + (\delta - 1)k \leqslant \lambda\rho/\rho_1 + 4n(\rho/\rho_1 - 1)/\rho.$$

Case 2. Now let $k \geqslant 4n/\rho$. The matrix HAH has A' at the upper right block and zeros everywhere else. According to Claim 13, we have $\operatorname{rank}(PA'Q) \leqslant n/k$ for any choice of p, q and P, Q as above. Using Claim 12, we find a $\mu \in \mathbb{F}$ for which the matrix $H_1 := HAH - \mu H$ satisfies $\rho_1 := \operatorname{rank}(H_1) \leqslant 2n/k$. So we have $\rho_1 \leqslant 0.5\rho$, and H_1 is spanned by words of length at most

$$2\lambda + k \leqslant \lambda\rho/\rho_1 + 2n/\rho_1 \leqslant \lambda\rho/\rho_1 + 4n(1 - \rho_1/\rho)/\rho_1. \qquad \square$$

Proof of Theorem 3. As in the proof of Theorem 9, we can assume without loss of generality that $\mathbb{F}$ is algebraically closed and S is irreducible. Using Claim 11, we find a square-zero matrix of rank $\rho_0 > 0$ in $\mathbb{F}S^{\leqslant\lambda_0}$ with $\lambda_0\rho_0 \leqslant 2n$; if $\rho_0 = 1$, then we apply Corollary 7 and complete the proof. Otherwise, we repeatedly apply Claim 14 and obtain a sequence $(\lambda_0, \rho_0), \dots, (\lambda_\tau, \rho_\tau)$ such that $\rho_\tau = 1$ and for all $t \in \{0, \dots, \tau - 1\}$ it holds that $\rho_{t+1} \in [1, 0.5\rho_t]$,

$$\lambda_{t+1} \leqslant \frac{\lambda_t\rho_t}{\rho_{t+1}} + \frac{4n(\rho_t - \rho_{t+1})}{\rho_t\rho_{t+1}},$$

and every $\mathbb{F}S^{\leqslant\lambda_t}$ contains a square-zero matrix of rank ρ_t. By induction we get

$$\lambda_t \leqslant \frac{\lambda_0\rho_0}{\rho_t} + \frac{4n}{\rho_t}\left(t - \frac{\rho_1}{\rho_0} - \cdots - \frac{\rho_t}{\rho_{t-1}}\right),$$

which implies (after the substitution $\alpha_t := \rho_t/\rho_{t-1}$) that

$$\lambda_\tau \leqslant 2n + 4n\left(\tau - \sum_{t=1}^{\tau} \alpha_t\right),$$

and since the minimum value of $\alpha_1 + \cdots + \alpha_\tau$ subject to $\alpha_t > 0$ and $\alpha_1 \cdots \alpha_\tau = \rho_0^{-1}$ is attained when $\alpha_1 = \cdots = \alpha_\tau = \rho_0^{-1/\tau}$, we get

$$\lambda_\tau \leqslant 2n + 4n\tau(1 - \rho_0^{-1/\tau}).$$

The right-hand side of this inequality is an increasing function of τ, so it attains its maximum at the largest possible value $\tau = \log_2 \rho_0$. We get $\lambda_\tau \leqslant 2n + 2n \log_2 \rho_0$, and it remains to apply Corollary 7. □

The author does not expect his result to be tight even asymptotically, so this paper does not show any effort on improving the $o(n \log n)$ part of the upper bound.

Acknowledgments

The author is indebted to O.V. Markova from Moscow State University for a series of talks on the topic, which he has had a privilege to attend since 2006. Mateusz Michałek told the author in June 2018 about a very similar problem, known as the quantum version of Wielandt's inequality [Sanz et al. 2010], and we quickly came to a conclusion that the progress on one of these problems can lead to the progress on the other. In particular, we managed [Michałek and Shitov 2019] to get an $O(n^2 \log n)$ bound for the largest value of τ such that the equality $\mathbb{F}S^\tau = \mathrm{Mat}_n(\mathbb{F})$ holds for any set S for which there exists a t satisfying $\mathbb{F}S^t = \mathrm{Mat}_n(\mathbb{F})$ — while the present paper gives an $O(n \log n)$ bound for the same problem but with S^τ, S^t replaced by $S^{\leqslant \tau}$, $S^{\leqslant t}$, respectively.

References

[Guterman et al. 2018] A. Guterman, T. Laffey, O. Markova, and H. Šmigoc, "A resolution of Paz's conjecture in the presence of a nonderogatory matrix", *Linear Algebra Appl.* **543** (2018), 234–250. MR Zbl

[Laffey 1986] T. J. Laffey, "Simultaneous reduction of sets of matrices under similarity", *Linear Algebra Appl.* **84** (1986), 123–138. MR Zbl

[Lambrou and Longstaff 2009] M. S. Lambrou and W. E. Longstaff, "On the lengths of pairs of complex matrices of size six", *Bull. Aust. Math. Soc.* **80**:2 (2009), 177–201. MR Zbl

[Longstaff et al. 2006] W. E. Longstaff, A. C. Niemeyer, and O. Panaia, "On the lengths of pairs of complex matrices of size at most five", *Bull. Aust. Math. Soc.* **73**:3 (2006), 461–472. MR Zbl

[Markova 2005] O. V. Markova, "On the length of the algebra of upper-triangular matrices", *Uspekhi Mat. Nauk* **60**:5 (2005), 177–178. In Russian; translated in *Russ. Math. Surv.* **60**:5 (2005), 984–985. MR Zbl

[Michałek and Shitov 2019] M. Michałek and Y. Shitov, "Quantum version of Wielandt's inequality revisited", *IEEE Trans. Inform. Theory* (online publication February 2019).

[Pappacena 1997] C. J. Pappacena, "An upper bound for the length of a finite-dimensional algebra", *J. Algebra* **197**:2 (1997), 535–545. MR Zbl

[Pappacena et al. 2003] C. J. Pappacena, L. W. Small, and J. Wald, "Affine semiprime algebras of GK dimension one are (still) pi", *Glasg. Math. J.* **45**:2 (2003), 243–247. MR Zbl

[Paz 1984] A. Paz, "An application of the Cayley–Hamilton theorem to matrix polynomials in several variables", *Linear Multilin. Algebra* **15**:2 (1984), 161–170. MR Zbl

[Radjavi and Rosenthal 2000] H. Radjavi and P. Rosenthal, *Simultaneous triangularization*, Springer, 2000. MR Zbl

[Sanz et al. 2010] M. Sanz, D. Pérez-García, M. M. Wolf, and J. I. Cirac, "A quantum version of Wielandt's inequality", *IEEE Trans. Inform. Theory* **56**:9 (2010), 4668–4673. MR Zbl

Communicated by Victor Reiner
Received 2018-11-14 Revised 2019-03-06 Accepted 2019-05-14

yaroslav-shitov@yandex.ru Moscow, Russia

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 13 No. 6 2019