Algebra & Number Theory

Volume 13
2019
No. 4

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/

msp

ALGEBRA AND NUMBER THEORY 13:4 (2019)
dx.doi.org/10.2140/ant.2019.13.749

Artin's criteria for algebraicity revisited

Jack Hall and David Rydh

Using notions of homogeneity we give new proofs of M. Artin's algebraicity criteria for functors and groupoids. Our methods give a more general result, unifying Artin's two theorems and clarifying their differences.

Introduction

Classically, moduli spaces in algebraic geometry are constructed using either projective methods or by forming suitable quotients. In his reshaping of the foundations of algebraic geometry half a century ago, Grothendieck shifted focus to the functor of points and the central question became whether certain functors are representable. Early on, he developed formal geometry and deformation theory, with the intent of using these as the main tools for proving representability. Grothendieck's proof of the existence of Hilbert and Picard schemes, however, is based on projective methods. It was not until ten years later that Artin completed Grothendieck's vision in a series of landmark papers. In particular, Artin vastly generalized Grothendieck's existence result and showed that the Hilbert and Picard schemes exist — as algebraic spaces — in great generality. It also became clear that the correct setting was that of algebraic spaces — not schemes — and algebraic stacks.

M. Artin [1969b; 1974] gave precise criteria for algebraicity of functors and stacks. These criteria were later clarified by B. Conrad and J. de Jong [2002] using Néron–Popescu desingularization, by H. Flenner [1981] using Exal, and the first author [Hall 2017] using coherent functors. The criterion in this last paper is very streamlined and elegant and suffices to deal with most problems. It does not, however, supersede Artin's criteria as these are more general. Another conundrum is that Artin gives two different criteria — one for functors in [Artin 1969b, Theorem 5.3] and one for stacks in [Artin 1974, Theorem 5.3] — but neither completely generalizes the other.

The purpose of this paper is to use the ideas of Flenner and the first author to give a new criterion that supersedes all present criteria. We also introduce several new ideas that broaden the criteria and simplify the proofs of [Artin 1969b; 1974; Flenner 1981]. In positive characteristic, we also identify a subtle issue in Artin's algebraicity criterion for stacks. With the techniques that we develop, this problem is circumvented. We now state our criterion for algebraicity.

This collaboration was supported by the Göran Gustafsson foundation. The first author was supported by the Australian Research Council DE150101799. The second author is also supported by the Swedish Research Council 2011-5599 and 2015-05554.
MSC2010: primary 14D15; secondary 14D23.
Keywords: algebraic stacks, deformation theory, obstruction theories.

Main Theorem. *Let S be an excellent scheme. Then a category X, fibered in groupoids over the category of S-schemes, Sch/S, is an algebraic stack, locally of finite presentation over S, if and only if it satisfies the following conditions*:

(1) *X is a stack over $(\mathsf{Sch}/S)_{\mathrm{fppf}}$.*

(2) *X is limit preserving* (Definition 1.7).

(3) *X is weakly effective* (Definition 9.1).

(4) *X is $\mathbf{Art}^{\mathrm{triv}}$-homogeneous* (Definition 1.3, *also see below*).

(5a) *X has bounded automorphisms and deformations* (*Conditions* 6.1(i)–(ii)).

(5b) *X has constructible automorphisms and deformations* (*Conditions* 6.3(i)–(ii)).

(5c) *X has Zariski local automorphisms and deformations* (*Conditions* 6.4(i)–(ii)).

(6b) *X has constructible obstructions* (*Condition* 6.3(iii), or Condition 7.3).

(6c) *X has Zariski local obstructions* (*Condition* 6.4(iii), or Condition 7.4).

In addition:

(α) *If S is Jacobson, then conditions (5c) and (6c) are superfluous.*

(β) *If X is* **DVR***-homogeneous* (Notation 2.14), *then conditions (5c) and (6c) are superfluous and condition (6b) may be replaced with* Condition 8.3.

(γ) *Conditions (1) and (4) can be replaced with these*:

($1'$) *X is a stack over $(\mathsf{Sch}/S)_{\acute{E}t}$.*

($4'$) *X is $\mathbf{Art}^{\mathrm{insep}}$-homogeneous.*

(δ) *If the residue fields of S at points of finite type are perfect, then (4) and ($4'$) are equivalent.*

In particular, if S is a scheme of finite type over $\operatorname{Spec}\mathbb{Z}$, *then conditions (5c) and (6c) are superfluous and (1) can be replaced with ($1'$).*

The $\mathbf{Art}^{\mathrm{triv}}$-homogeneity (resp. $\mathbf{Art}^{\mathrm{insep}}$-homogeneity) condition is the following Schlessinger–Rim condition: for every diagram of local artinian S-schemes of finite type $[\operatorname{Spec} B \leftarrow \operatorname{Spec} A \hookrightarrow \operatorname{Spec} A']$, where $A' \twoheadrightarrow A$ is surjective and the residue field extension $B/\mathfrak{m}_B \to A/\mathfrak{m}_A$ is trivial (resp. purely inseparable), the natural functor

$$X(\operatorname{Spec}(A' \times_A B)) \to X(\operatorname{Spec} A') \times_{X(\operatorname{Spec} A)} X(\operatorname{Spec} B)$$

is an equivalence of categories.

Perhaps the most striking difference between our conditions and Artin's conditions is that our homogeneity condition (4) only involves local artinian schemes and that we do not need any conditions on étale localization of deformation and obstruction theories. If S is Jacobson, e.g., of finite type over a field, then we do not even need compatibility with Zariski localization. There is also no condition on compatibility

with completions for automorphisms and deformations. We will give a detailed comparison between our conditions and other versions of Artin's conditions in Section 11.

All existing algebraicity proofs, including ours, consist of the following four steps:

(i) existence of formally versal deformations;

(ii) algebraization of formally versal deformations;

(iii) openness of formal versality; and

(iv) formal versality implies formal smoothness.

Step (i) was eloquently dealt with by Schlessinger [1968, Theorem 2.11] for functors and by Rim [SGA 7_{I} 1972, Exposé VI] for groupoids. This step uses conditions (4) and (5a) ($\mathbf{Art}^{\mathrm{triv}}$-homogeneity and boundedness of tangent spaces). Step (ii) begins with the effectivization of formally versal deformations using condition (3). One may then algebraize this family using either Artin's results [1969a; 1969b] or B. Conrad and J. de Jong's result [2002]. In the latter approach, Artin approximation is replaced with Néron–Popescu desingularization, and S is only required to be excellent. This step requires condition (2).

The last two steps are more subtle and it is here that [Artin 1969b; 1974; Flenner 1981; Starr 2006; Hall 2017] and our present treatment diverge — both when it comes to the criteria themselves and the techniques employed. We begin with discussing step (iv).

Formal versality implies formal smoothness. It is readily seen that our criterion is weaker than Artin's two criteria [1969b; 1974] except that, in positive characteristic, we need X to be a stack in the fppf topology, or otherwise strengthen (4). This is similar to [Artin 1969b, Theorem 5.3] where the functor is assumed to be an fppf-sheaf. Artin [1969b, Theorem 5.3] deftly uses the fppf sheaf condition to deduce that formally universal deformations are formally étale [loc. cit., pp. 50–52], settling step (iv) for functors. This argument relies on the existence of universal deformations and thus does not extend to stacks with infinite or nonreduced stabilizers. Using a different approach, we extend this result to fppf stacks in Lemma 1.9.

In his second paper, Artin [1974] only assumes that the groupoid is an étale stack. His proof of step (iv) for groupoids [loc. cit., Proposition 4.2], however, does not treat inseparable extensions. We do not understand how this problem can be overcome without strengthening the criteria and assuming that either (1) the groupoid is a stack in the fppf topology or (4′) requiring (semi)homogeneity for inseparable extensions (see Lemmas 1.9 and 2.2). We wish to emphasize that if S is of finite type over $\operatorname{Spec}\mathbb{Z}$ or a perfect field, then the main result of [loc. cit.] holds without change. See Remark 2.8 for further discussion. Flenner does not discuss formal smoothness, and in [Hall 2017] formal smoothness is obtained by strengthening the homogeneity condition (4).

Openness of formal versality. Step (iii) uses constructibility, boundedness, and Zariski localization of deformations and obstruction theories (Theorem 4.4). In our treatment, localization is only required when passing to nonclosed points of finite type. Such points only exist when S is not Jacobson, e.g., if S is the

spectrum of a discrete valuation ring. Our proof is very similar to Flenner's proof. It may appear that Flenner does not need Zariski localization in his criterion, but this is due to the fact that his conditions are expressed in terms of deformation and obstruction *sheaves*.

As in Flenner's proof, openness of versality becomes a matter of simple algebra. It comes down to a criterion for the openness of the *vanishing locus* of half-exact functors (Theorem 3.3) that easily follows from the Ogus–Bergman Nakayama Lemma for half-exact functors (Theorem 3.7). Flenner proves a stronger statement that implies the Ogus–Bergman result (Remark 3.8).

At first, it seems that we need more than $\mathbf{Art}^{\mathrm{triv}}$-homogeneity to even make sense of conditions (5a)–(6c). This will turn out to not be the case. Using steps (ii) and (iv), we prove that conditions (1)–(4) and (5a) at fields guarantee that we have homogeneity for arbitrary integral morphisms (Lemma 10.4). It follows that $\mathrm{Aut}_{X/S}(T, -)$, $\mathrm{Def}_{X/S}(T, -)$ and $\mathrm{Obs}_{X/S}(T, -)$ are additive functors.

Applications. We believe that a distinct advantage of the criterion in the present paper contrasted with all prior criteria is the dramatic weakening of the homogeneity. Whereas the criteria [Hall 2017; Artin 1969b] require **Aff**-, and **DVR**-homogeneity respectively, involving knowledge of the functor over nonnoetherian rings, we only need homogeneity for artinian rings. This is particularly useful for more subtle moduli problems such as Angéniol's Chow functor [1981, Théorème 5.2.1], which is difficult to define over nonnoetherian rings.

The ideas in this paper have also led to a criterion for a half-exact functor to be coherent [Hall and Rydh 2013]. Although both the statement and the proof bear a close resemblance to the Main Theorem, this coherence criterion does not follow from any algebraicity criterion.

Outline. The technical results of the paper are summarized by Proposition 10.2. The Main Theorem follows from Proposition 10.2 by a bootstrapping process and the relationship between automorphisms, deformations, obstructions and extensions. A significant part of the paper (Sections 5–9) is devoted to making this relationship precise. Sections 1–4 form the technical heart of the paper. We now briefly summarize the contents of the paper in more detail.

In Section 1 we recall the notions of homogeneity, limit preservation and extensions from [Hall 2017]. We also introduce homogeneity that only involves artinian rings and show that residue field extensions are harmless for stacks in the fppf topology. In Section 2 we then relate formal versality, formal smoothness and vanishing of Exal.

In Section 3 we study additive functors and their vanishing loci. This is applied in Section 4 where we give conditions on Exal that assure that the locus of formal versality is open. The results are then assembled in Theorem 4.4.

In Section 5 we repeat the definitions of automorphisms, deformations and minimal obstruction theories from [Hall 2017]. In Section 6, we give conditions on Aut, Def and Obs that imply the corresponding conditions on Exal needed in Theorem 4.4. In Section 7 we introduce n-step obstruction theories. In Section 8 we formulate the conditions on obstructions without using linear obstruction theories, as

in [Artin 1969b]. In Section 9, we discuss effectivity. Finally, in Section 10 we prove the Main Theorem. Comparisons with other criteria are given in Section 11.

Notation. We follow standard conventions and notation. In particular, we adhere to the notation of [Hall 2017]. Recall that if T is a scheme, then a point $t \in |T|$ is of *finite type* if $\operatorname{Spec} \kappa(t) \to T$ is of finite type. Points of finite type are locally closed. A point of a Jacobson scheme is of finite type if and only if it is closed. If $f\colon X \to Y$ is of finite type and $x \in |X|$ is of finite type, then $f(x) \in |Y|$ is of finite type.

1. Homogeneity, limit preservation, and extensions

Fix a scheme S. An *S-groupoid* is a category X together with a functor $a_X\colon X \to \mathsf{Sch}/S$ that is fibered in groupoids. A 1-morphism of S-groupoids $\Phi\colon (Y, a_Y) \to (Z, a_Z)$ is a functor between categories Y and Z that commutes strictly over Sch/S. We will typically refer to an S-groupoid (X, a_X) as "X".

A closed immersion of schemes $j\colon V \hookrightarrow V'$ is *nilpotent* if there exists an integer $n > 0$ such that $J^n = 0$, where J is the quasicoherent sheaf of ideals defining j. A closed immersion of schemes is *locally nilpotent* if fppf-locally it is nilpotent.

If X is an S-groupoid and $[\operatorname{Spec} B \leftarrow \operatorname{Spec} A \xrightarrow{j} \operatorname{Spec} A']$ is a diagram of S-schemes, where j is a nilpotent closed immersion, then the condition that the functor

$$X(\operatorname{Spec}(B \times_A A')) \to X(\operatorname{Spec} B) \times_{X(\operatorname{Spec} A)} X(\operatorname{Spec} A')$$

is an equivalence for a collection of diagrams has been a feature of deformation theory since Schlessinger [1968] and Rim [SGA 7_{I} 1972, Exposé VI]. Consequently, these are typically called *Schlessinger–Rim* conditions.

In this section, we review the concept of homogeneity — a variation of the Schlessinger–Rim conditions that we attribute to J. Wise [2011, §2] — in the formalism of [Hall 2017, §1–2]. We will also briefly discuss limit preservation and extensions.

Let X be an S-groupoid. An *X-scheme* is a pair (T, σ_T), where T is an S-scheme and $\sigma_T\colon \mathsf{Sch}/T \to X$ is a 1-morphism of S-groupoids. A morphism of X-schemes $U \to V$ is a morphism of S-schemes $f\colon U \to V$ (which canonically determines a 1-morphism of S-groupoids $\mathsf{Sch}/f\colon \mathsf{Sch}/U \to \mathsf{Sch}/V$) together with a 2-morphism $\alpha\colon \sigma_U \Rightarrow \sigma_V \circ \mathsf{Sch}/f$. The collection of all X-schemes forms a 1-category, which we denote by Sch/X. It is readily seen that Sch/X is an S-groupoid and that there is a natural equivalence of S-groupoids $\mathsf{Sch}/X \to X$. For a 1-morphism of S-groupoids $\Phi\colon Y \to Z$ there is an induced functor $\mathsf{Sch}/\Phi\colon \mathsf{Sch}/Y \to \mathsf{Sch}/Z$.

Notation 1.1. Frequently, we will be interested in the following classes of morphisms of S-schemes:

Nil: locally nilpotent closed immersions,

Cl: closed immersions,

rNil: morphisms $X \to Y$ such that there exists $(X_0 \to X) \in \mathbf{Nil}$ with the composition $(X_0 \to X \to Y) \in \mathbf{Nil}$,

rCl: morphisms $X \to Y$ such that there exists $(X_0 \to X) \in \mathbf{Nil}$ with the composition $(X_0 \to X \to Y) \in \mathbf{Cl}$,

$\mathbf{Art}^{\mathrm{fin}}$: morphisms between local artinian schemes of finite type over S,

$\mathbf{Art}^{\mathrm{sep}}$: $\mathbf{Art}^{\mathrm{fin}}$-morphisms with separable residue field extensions,

$\mathbf{Art}^{\mathrm{insep}}$: $\mathbf{Art}^{\mathrm{fin}}$-morphisms with purely inseparable residue field extensions,

$\mathbf{Art}^{\mathrm{triv}}$: $\mathbf{Art}^{\mathrm{fin}}$-morphisms with trivial residue field extensions,

Fin: finite morphisms,

Int: integral morphisms, and

Aff: affine morphisms.

We certainly have a containment of classes of morphisms of S-schemes:

$$\begin{array}{ccccccc} & & \mathbf{rNil} & \subseteq & \mathbf{rCl} & \subseteq & \mathbf{Int} & \subseteq & \mathbf{Aff}. \\ & \diagup\!\diagup & \cup\mathrm{I} & & \cup\mathrm{I} & & \cup\mathrm{I} & & \\ & & \mathbf{Nil} & \subseteq & \mathbf{Cl} & \subseteq & \mathbf{Fin} & & \\ & & & & & & \cup\mathrm{I} & & \\ \mathbf{Art}^{\mathrm{triv}} & \subseteq & \mathbf{Art}^{\mathrm{insep}} & & \subseteq & & \mathbf{Art}^{\mathrm{fin}} & & \end{array}$$

Note that for a morphism $X \to Y$ of locally noetherian S-schemes, the properties **rNil** and **rCl** simply mean that $X_{\mathrm{red}} \to Y$ is **Nil** and **Cl** respectively. The classes of morphisms above are all closed under composition.

Let P be a class of morphisms of S-schemes. In [Hall 2017, §1], P-nil pairs and P-homogeneity were defined. In the present article, it will be necessary to consider some natural refinements of these notions.

Definition 1.2. Fix a scheme S, a class P of morphisms of S-schemes, an S-groupoid X and an X-scheme V. A *P-nil pair over X at V* is a pair $(V \xrightarrow{p} T, V \xrightarrow{j} V')$, where p and j are morphisms of X-schemes, $p \in P$ and $j \in \mathbf{Nil}$. A *P-nil square over X at V* is a commutative diagram of X-schemes

$$\begin{array}{ccc} V & \xrightarrow{p} & T \\ {\scriptstyle j}\downarrow & & \downarrow{\scriptstyle i} \\ V' & \xrightarrow{p'} & T' \end{array} \tag{1-1}$$

where the pair $(V \xrightarrow{p} T, V \xrightarrow{j} V')$ is P-nil over X at V. A P-nil square over X at V is *cocartesian* if it is cocartesian in the category of X-schemes. A P-nil square over X at V is *geometric* if p' is affine, i is a locally nilpotent closed immersion, and there is a natural isomorphism

$$\mathcal{O}_{T'} \to i_*\mathcal{O}_T \times_{p'_* j_* \mathcal{O}_V} p'_* \mathcal{O}_{V'}.$$

Note that every geometric P-nil square is cartesian [Ferrand 2003, Lemme 1.3c]. Moreover if $P \subseteq \mathbf{Aff}$, then every cocartesian P-nil square is geometric [Hall 2017, Lemma 1.5(1)].

Definition 1.3 (P-homogeneity). Fix a scheme S and a class P of morphisms of S-schemes. A 1-morphism of S-groupoids $\Phi\colon Y \to Z$ is *P-homogeneous at a Y-scheme V* if the following two conditions are satisfied:

($^V\mathrm{H}_1^P$) A P-nil square over Y at V is cocartesian if and only if the induced P-nil square over Z at V is cocartesian.

($^V\mathrm{H}_2^P$) If a P-nil pair over Y at V can be completed to a cocartesian P-nil square over Z at V, then it can be completed to a P-nil square over Y at V.

We also say that Φ is *P-homogeneous* if it is P-homogeneous at every Y-scheme V. Similarly, Φ satisfies (H_1^P) (resp. (H_2^P)) if it satisfies ($^V\mathrm{H}_1^P$) (resp. ($^V\mathrm{H}_2^P$)) for every Y-scheme V. An S-groupoid X is *P-homogeneous at V* if its structure 1-morphism is P-homogeneous at V and is *P-homogeneous* if its structure morphism is P-homogeneous. If Z satisfies (H_1^P), then Y satisfies (H_1^P) if and only if Φ has P-homogeneous diagonal after pull-back to schemes, see Lemma B.2.

If we only assume ($^V\mathrm{H}_2^P$) in the above, then we obtain the weaker notion of *P-semihomogeneity*. This notion was used in the work of Artin and Flenner.

Remark 1.4. In [Hall 2017], a number of results are established for 1-morphisms of P-homogeneous S-groupoids $\Phi\colon Y \to Z$. With trivial modifications, most of these results hold using the more refined notion of P-homogeneity at a Y-scheme V. We will use this observation frequently and without further comment.

By [Wise 2011, Propostion 2.1], every algebraic stack is **Aff**-homogeneous. Also, **rNil**-homogeneity at an artinian scheme V is equivalent to $\mathbf{Art}^{\mathrm{triv}}$-homogeneity at V.

If P is Zariski local (e.g., P is listed in Notation 1.1), then P-homogeneity of an S-groupoid X that is a stack over $(\mathsf{Sch}/S)_{\text{Ét}}$ is equivalent to the functor

$$X(\operatorname{Spec}(B \times_A A')) \to X(\operatorname{Spec} B) \times_{X(\operatorname{Spec} A)} X(\operatorname{Spec} A') \tag{1-2}$$

being an equivalence for every P-nil pair ($\operatorname{Spec} A \to \operatorname{Spec} B$, $\operatorname{Spec} A \to \operatorname{Spec} A'$) over S [Hall 2017, Lemma 1.5(4)]. If X has representable diagonal, then the functor above is always fully faithful for all **Aff**-nil pairs over S — even if X is not necessarily **Aff**-homogeneous (Lemma B.2).

The main computational tools that P-homogeneity bring are contained in [Hall 2017, Lemma 1.5], an important part of which we now recall.

Lemma 1.5. *Let S be a scheme and let $P \subseteq \mathbf{Aff}$ be a class of morphisms. Let X be an S-groupoid that is P-homogeneous at an X-scheme V. If $(V \xrightarrow{p} T, V \xrightarrow{j} V')$ is a P-nil pair at V, then there exists a cocartesian and geometric P-nil square at V as in (1-1). Moreover if P is listed in Notation 1.1, then p' is P.*

Proof. The main claim is [Hall 2017, Lemma 1.5(3)]. What remains is trivial except for $P \in \{\mathbf{Nil}, \mathbf{Cl}, \mathbf{Fin}, \mathbf{Int}\}$. In these cases, however, it is known [Ferrand 2003, Proposition 5.6 (3)]. □

Remark 1.6. Let S be a noetherian scheme. If $(\operatorname{Spec} A \to \operatorname{Spec} B, \operatorname{Spec} A \hookrightarrow \operatorname{Spec} A']$ is a **Fin**-nil pair, where $\operatorname{Spec} B$ is of finite type over S, then $\operatorname{Spec}(B \times_A A')$ is of finite type over S. This follows from the fact that $B \times_A A' \subseteq B \times A'$ is an integral extension [Atiyah and Macdonald 1969, Proposition 7.8]. On the other hand, if $\operatorname{Spec} A \to \operatorname{Spec} B$ is only affine, then $\operatorname{Spec}(B \times_A A')$ is typically not of finite type over S. For example, if $B = k[x]$, $A = k[x, x^{-1}]$ and $A' = k[x, x^{-1}, y]/y^2$, then $B' = B \times_A A' = k[x, y, yx^{-1}, yx^{-2}, \dots]/(y, yx^{-1}, \dots)^2$ which is not of finite type over $S = \operatorname{Spec} k$.

We also recall the following definition (see [Artin 1974, §1; Hall 2017, §3]).

Definition 1.7. Let X be a stack over $(\mathsf{Sch}/S)_{\text{Ét}}$. We say that X is *limit preserving* if for every inverse system of affine S-schemes $\{\operatorname{Spec} A_j\}_{j\in J}$ with inverse limit $\operatorname{Spec} A$, the natural functor:

$$\varinjlim_j X(\operatorname{Spec} A_j) \to X(\operatorname{Spec} A)$$

is an equivalence of categories.

If X is an algebraic stack, then X is limit preserving if and only if $X \to S$ is locally of finite presentation [Laumon and Moret-Bailly 2000, Proposition 4.15].

By Lemmas B.2 and B.3, if X is a limit preserving stack over $(\mathsf{Sch}/S)_{\text{Ét}}$ with representable diagonal and S is locally noetherian, then **rCl**-homogeneity is equivalent to Artin's *semihomogeneity* condition [1974, 2.2(S1a)] for X.

Homogeneity supplies an S-groupoid with a quantity of linear data, which we now recall from [Hall 2017, §2]. An *X-extension* is a square zero closed immersion of X-schemes $i\colon T \hookrightarrow T'$. The collection of X-extensions forms a category, which we denote by $\mathbf{Exal}_X$. There is a natural functor $\mathbf{Exal}_X \to \mathsf{Sch}/X$ that takes $(i\colon T \hookrightarrow T')$ to T.

We denote by $\mathbf{Exal}_X(T)$ the fiber of the category $\mathbf{Exal}_X$ over the X-scheme T — we call these the X-extensions of T. There is a natural functor:

$$\mathbf{Exal}_X(T)^\circ \to \mathsf{QCoh}(T), \quad (i\colon T \hookrightarrow T') \mapsto \ker(i^{-1}\mathcal{O}_{T'} \to \mathcal{O}_T).$$

We denote by $\mathbf{Exal}_X(T, I)$ the fiber category of $\mathbf{Exal}_X(T)$ over the quasicoherent $\mathcal{O}_T$-module I — we refer to these as the X-extensions of T by I. Denote the set of isomorphism classes of the category $\mathbf{Exal}_X(T, I)$ by $\operatorname{Exal}_X(T, I)$.

Let W be a scheme and let J be a quasicoherent $\mathcal{O}_W$-module. We let $W[J]$ denote the W-scheme $\underline{\operatorname{Spec}}_W(\mathcal{O}_W[J])$ with structure morphism $r_{W,J}\colon W[J] \to W$. If W is an X-scheme, we consider $W[J]$ as an X-scheme via $r_{W,J}$. The X-extension $W \hookrightarrow W[J]$ is thus *trivial* in the sense that it admits an X-retraction.

By [Hall 2017, Proposition 2.4], if the S-groupoid X is **Nil**-homogeneous at T, then the groupoid $\mathbf{Exal}_X(T, I)$ is a Picard category. Thus, we have additive functors

$$\operatorname{Der}_X(T, -)\colon \mathsf{QCoh}(T) \to \mathsf{Ab}, \quad I \mapsto \operatorname{Aut}_{\mathbf{Exal}_X(T,I)}(T[I]); \quad \text{and}$$
$$\operatorname{Exal}_X(T, -)\colon \mathsf{QCoh}(T) \to \mathsf{Ab}, \quad I \mapsto \operatorname{Exal}_X(T, I).$$

We now record here the following easy consequences of [Hall 2017, 2.3–2.6, 3.4].

Lemma 1.8. *Let S be a scheme, let X be an S-groupoid, and let T be an X-scheme.*

(1) *Let I be a quasicoherent $\mathcal{O}_T$-module. Then $\mathrm{Exal}_X(T, I) = 0$ if and only if every X-extension $i : T \hookrightarrow T'$ of T by I admits an X-retraction.*

(2) *Let P be a class of a morphisms of S-schemes and let $p : V \to T$ be an affine morphism in P. If X is P-homogeneous at V, then for every $N \in \mathsf{QCoh}(V)$ there is a natural functor*

$$p_{\#} : \mathbf{Exal}_X(V, N) \to \mathbf{Exal}_X(T, p_*N).$$

(3) *If X is* **rNil***-homogeneous at T, then the functor $M \mapsto \mathrm{Exal}_X(T, M)$ is half-exact.*

(4) *Suppose that X is* **Nil***-homogeneous at T and limit preserving. If T is of finite presentation over S, then the functor $M \mapsto \mathrm{Exal}_X(T, M)$ preserves direct limits.*

(5) *Let $p : U \to T$ be an* ***affine*** *étale morphism and let N be a quasicoherent $\mathcal{O}_U$-module. Then there is a natural functor $\psi : \mathbf{Exal}_X(T, p_*N) \to \mathbf{Exal}_X(U, N)$. If $(i : T \hookrightarrow T') \in \mathbf{Exal}_X(T, p_*N)$ with image $(j : U \hookrightarrow U') \in \mathbf{Exal}_X(U, N)$, then there is a cartesian diagram of X-schemes*

$$\begin{array}{ccc} U & \xhookrightarrow{j} & U' \\ {\scriptstyle p}\downarrow & & \downarrow{\scriptstyle p'} \\ T & \xhookrightarrow{i} & T' \end{array}$$

which is cocartesian as a diagram of S-schemes. If X is **Aff***-homogeneous at U, then ψ is an equivalence.*

Proof. The claim (1) is [Hall 2017, Lemma 2.3].

For (2), if $j : V \hookrightarrow V'$ is an X-extension of V by N, then there is an induced P-nil pair $(V \xrightarrow{p} T, V \xrightarrow{j} V')$ over X at V. Since X is P-homogeneous at V, by Lemma 1.5, there exists a cocartesian and geometric P-nil square over X at V as in (1-1) completing the P-nil pair. The resulting morphism $i : T \hookrightarrow T'$ is an X-extension of T by p_*N and we have thus defined the functor $p_{\#}$.

The claim (3) is [Hall 2017, Corollary 2.5]. The claim (4) is [loc. cit., Proposition 3.4(2)]. The claim (5) is [loc. cit., Corollary 2.6]. □

Finally, we give conditions that imply $\mathbf{Art}^{\mathrm{sep}}$- and $\mathbf{Art}^{\mathrm{fin}}$-homogeneity.

Lemma 1.9. *Let S be a scheme and let X be an S-groupoid that is $\mathbf{Art}^{\mathrm{triv}}$-homogeneous. Consider the following conditions on X:*

(1) *X is a stack in the fppf topology.*

(2) *X is a stack in the étale topology and $\mathbf{Art}^{\mathrm{insep}}$-homogeneous.*

(3) *X is a stack in the étale topology and S is a $\mathbb{Q}$-scheme.*

(4) *X is a stack in the étale topology.*

Then any of the conditions (1), (2), *or* (3) *imply that* X *is* $\mathbf{Art}^{\mathrm{fin}}$*-homogeneous and condition* (4) *implies that* X *is* $\mathbf{Art}^{\mathrm{sep}}$*-homogeneous.*

Proof. We begin by noting that trivially (3) implies (2). Next, let $(\operatorname{Spec} A \to \operatorname{Spec} B, \operatorname{Spec} A \hookrightarrow \operatorname{Spec} A')$ be an $\mathbf{Art}^{\mathrm{fin}}$-nil pair over S. Let $\operatorname{Spec} B' = \operatorname{Spec}(A' \times_A B)$ be the pushout of this diagram in the category of S-schemes. We have to prove that the functor

$$\varphi\colon X(\operatorname{Spec} B') \to X(\operatorname{Spec} A') \times_{X(\operatorname{Spec} A)} X(\operatorname{Spec} B)$$

is an equivalence. If X is a stack in either the fppf or étale topology, then the equivalence of φ is a local question for the respective topology on B' since fiber products of rings commute with flat base change.

Now there is a finite (resp. finite separable) field extension K/k_B such that the residue fields of $k_A \otimes_{k_B} K$ are trivial (resp. purely inseparable) extensions of K. There is then a local artinian ring $\widetilde{B}'$ and a finite flat (resp. finite étale) extension $B' \hookrightarrow \widetilde{B}'$ with $k_{\widetilde{B}'} = K$ [EGA III$_1$ 1961, Corollaire 0.10.3.2]. Let $\widetilde{A} = A \otimes_{B'} \widetilde{B}'$, $\widetilde{A}' = A' \otimes_{B'} \widetilde{B}'$ and $\widetilde{B} = B \otimes_{B'} \widetilde{B}'$. Then $\widetilde{A}$, $\widetilde{A}'$, $\widetilde{B}$ are artinian rings such that all residue fields equal K (resp. are purely inseparable extensions of K). However, $\widetilde{A}$ and $\widetilde{A}'$ need not be local. Now let $\widetilde{A} = \prod_{i=1}^n \widetilde{A}_i$ and $\widetilde{A}' = \prod_{i=1}^n \widetilde{A}'_i$ be decompositions such that $\widetilde{A}' \twoheadrightarrow \widetilde{A}_i$ factors through $\widetilde{A}'_i$. Then $\widetilde{B}' = (\widetilde{A}'_1 \times_{\widetilde{A}_1} \widetilde{B}) \times_{\widetilde{B}} (\widetilde{A}'_2 \times_{\widetilde{A}_2} \widetilde{B}) \times_{\widetilde{B}} \cdots \times_{\widetilde{B}} (\widetilde{A}'_n \times_{\widetilde{A}_n} \widetilde{B})$ is an iterated fiber product of local artinian rings.

If X is $\mathbf{Art}^{\mathrm{triv}}$-homogeneous (resp. $\mathbf{Art}^{\mathrm{insep}}$-homogeneous) and a stack for the fppf (resp. étale) topology, it follows that φ is an equivalence. If the $\mathbf{Art}^{\mathrm{fin}}$-nil pair that we started with was an $\mathbf{Art}^{\mathrm{sep}}$-nil pair and X is a stack for the étale topology, then it also follows that φ is an equivalence. This proves the result. □

2. Formal versality and formal smoothness

In this section we address a subtle point about the relationship between formal versality and formal smoothness. We begin by recalling and refining some results of [Hall 2017, §4].

Definition 2.1. Let S be a scheme, let X be an S-groupoid, and let T be an X-scheme. Consider the following lifting problem in the category of X-schemes: given a pair of morphisms of X-schemes $(V \xrightarrow{p} T, V \xrightarrow{j} V')$, where j is a locally nilpotent closed immersion, complete the following diagram so that it commutes:

$$\begin{array}{ccc} V & \xrightarrow{p} & T \\ {\scriptstyle j}\downarrow & \nearrow & \\ V' & & \end{array} \tag{2-1}$$

The X-scheme T is

formally smooth if the lifting problem can always be solved Zariski-locally on V';

formally smooth at $t \in |T|$ if the lifting problem can always be solved whenever the X-schemes V and V' are local artinian, with closed points v and v', respectively, such that $p(v) = t$, and the field extension $\kappa(t) \subseteq \kappa(v)$ is finite;

formally versal at $t \in |T|$ if the lifting problem can always be solved whenever the X-schemes V and V' are local artinian, with closed points v and v', respectively, such that $p(v) = t$, and the field extension $\kappa(t) \subseteq \kappa(v)$ is an isomorphism.

We certainly have the following implications:

$$\text{formally smooth} \Longrightarrow \text{formally smooth at all } t \in |T| \Longrightarrow \text{formally versal at all } t \in |T|.$$

Formal smoothness and formal versality at all $t \in |T|$ are not obviously equivalent. Even for morphisms of finite type between noetherian schemes, it is a nontrivial result that they are equivalent [EGA IV$_4$ 1967, Proposition 17.14.2] (also see [Stacks Project, Tag 02HX] and Corollary 2.5).

Formal smoothness at t and formal versality at t are also not obviously equivalent. Moreover without stronger assumptions, it is not obvious to the authors that formal smoothness or formal versality is smooth-local on the source. We will see, however, that these subtleties vanish whenever the S-groupoid is $\mathbf{Art}^{\mathrm{fin}}$-homogeneous. For formal versality and formal smoothness at a point, it is sufficient that liftings exist when $\kappa(v) \cong j^{-1} \ker(\mathcal{O}_{V'} \to \mathcal{O}_V)$.

The goal of this section is to give sufficient conditions for a family, formally versal at all *closed* points, to be formally *smooth*. In Artin's papers, Artin approximation is used to address this. With our formulation, excellence (or related) assumptions are irrelevant. For some further discussion on Artin's approach, see Remark 2.8.

There is a tight connection between formal smoothness (resp. formal versality) and X-extensions in the *affine* setting. Most of the next result was proved in [Hall 2017, Lemma 4.3], which utilized arguments similar to those of [Flenner 1981, Satz 3.2].

Lemma 2.2. *Let S be a scheme, let X be an S-groupoid, and let T be an* ***affine*** *X-scheme. Let $t \in |T|$ be a point. Consider the following conditions*:

(1) *The X-scheme T is formally smooth at t.*

(2) *The X-scheme T is formally versal at t.*

(3) *X is* **Nil***-homogeneous at T and* $\mathrm{Exal}_X(T, \kappa(t)) = 0$.

Then (1) $\Longrightarrow$ (2) *and if X is* $\mathbf{Art}^{\mathrm{fin}}$*-semihomogeneous and t is of finite type, then* (2) $\Longrightarrow$ (1). *If X is* **Cl***-homogeneous, T is noetherian and t is a* ***closed*** *point, then* (2) $\Longrightarrow$ (3). *If X is* **rCl***-homogeneous and t is a* ***closed*** *point, then* (3) $\Longrightarrow$ (2).

Thus, assuming that an S-groupoid X is **rCl**-homogeneous, we can reformulate formal versality of an affine X-scheme T at a closed point $t \in |T|$ in terms of the triviality of the abelian group $\mathrm{Exal}_X(T, \kappa(t))$. Understanding the set of points $U \subseteq |T|$ where $\mathrm{Exal}_X(T, \kappa(u)) = 0$ for $u \in |U|$ will be accomplished in the next section.

Remark 2.3. If X is **Aff**-homogeneous and $\mathrm{Exal}_X(T, -) \equiv 0$, then T is formally smooth [Hall 2017, Lemma 4.3] but we will not use this. If Exal_X commutes with Zariski localization, that is, if for every open immersion of affine schemes $U \subseteq T$ the canonical map $\mathrm{Exal}_X(T, M) \otimes_{\Gamma(\mathcal{O}_T)} \Gamma(\mathcal{O}_U) \to \mathrm{Exal}_X(U, M|_U)$

is bijective, then the implications (2) $\implies$ (3) and (3) $\implies$ (2) also hold for nonclosed points. This is essentially what Flenner [1981, Satz 3.2] proves as his $\mathcal{E}x(T \to X, M)$ is the sheafification of the presheaf $U \mapsto \mathrm{Exal}_X(U, M|_U)$.

Proof of Lemma 2.2. The implication (1) $\implies$ (2) follows from the definition. The implications (2) $\implies$ (3) and (3) $\implies$ (2) are proved in [Hall 2017, Lemma 4.3]. The implication (2) $\implies$ (1) follows from a similar argument: assume that T is formally versal at t and fix a lifting problem as in diagram (2-1), where $j\colon V \to V'$ is a closed immersion of local artinian schemes with closed points v and v', respectively, such that $p(v) = t$ and $\kappa(v)/\kappa(t)$ is a finite extension. Let W be the schematic image of $V \to \mathrm{Spec}(\mathcal{O}_{T,t})$. Then W is a local artinian scheme with residue field $\kappa(t)$. As X is $\mathbf{Art}^{\mathrm{fin}}$-semihomogeneous, the $\mathbf{Art}^{\mathrm{fin}}$-nil pair $(V \to W, V \xrightarrow{j} V')$ over X can be completed to a geometric $\mathbf{Art}^{\mathrm{fin}}$-nil square over X:

$$\begin{array}{ccc} V & \longrightarrow & W \\ \downarrow & & \downarrow \\ V' & \longrightarrow & W' \end{array}$$

where $W \hookrightarrow W'$ is a closed immersion of local artinian schemes. Since the closed point of W has the same residue field as that of t, by formal versality, we obtain a lift of $W \to T$ to $W' \to T$ over X. The result follows. □

Lemma 2.2 is already quite powerful. In the following Proposition, we give a simple proof of [EGA IV$_1$ 1964, Proposition 0.22.1.4] in the case of a finitely generated or separable extension of residue fields (also see [Stacks Project, Tag 02HT]).

Proposition 2.4. *Let $f\colon T \to X$ be a morphism of locally noetherian schemes and let $t \in |T|$ with image $x = f(t)$. Consider the following conditions*:

(1) *The ring homomorphism $\mathcal{O}_{X,x} \to \mathcal{O}_{T,t}$ is preadically formally smooth* [EGA IV$_1$ 1964, Définition 0.19.3.1].

(2) *f is formally smooth at t.*

(3) *f is formally versal at t.*

Then (1) $\implies$ (2) $\iff$ (3). *If $\kappa(x) \subseteq \kappa(t)$ is finitely generated or separable, then* (3) $\implies$ (1).

Proof. We recall [EGA IV$_1$ 1964, Définition 0.19.3.1] for our situation. The preadic topology on a noetherian local ring has as a basis of open neighborhoods the powers of the maximal ideal. A local ring homomorphism $(A, \mathfrak{m}) \to (B, \mathfrak{n})$, where A and B are noetherian and preadically topologized, is smooth for the preadic topologies if for every discrete and continuous A-algebra C and nilpotent ideal $I \subseteq C$, all continuous A-algebra homomorphisms $B \to C/I$ factor continuously as $B \to C \to C/I$. Since A and B have their preadic topologies, this means that we can choose $n \gg 0$ such that $A \to C$ factors through $A \to A/\mathfrak{m}^n$ and $B \to C/I$ factors through $B \to B/\mathfrak{n}^n$. Note that both $A/\mathfrak{m}^n$ and $B/\mathfrak{n}^n$ are local artinian. Hence, (1) $\implies$ (2) $\implies$ (3).

For (3) $\Longrightarrow$ (2): we may assume that $X = \operatorname{Spec} \mathcal{O}_{X,x}$ and $T = \operatorname{Spec} \mathcal{O}_{T,t}$. In particular, $t \in |T|$ is a finite type point and X is $\mathbf{Art}^{\mathrm{fin}}$-homogeneous. By Lemma 2.2, the claim follows.

To prove (3) $\Longrightarrow$ (1) we will take $(A, \mathfrak{m}) = (\mathcal{O}_{X,x}, \mathfrak{m}_x)$ and $(B, \mathfrak{n}) = (\mathcal{O}_{T,t}, \mathfrak{m}_t)$ and consider the lifting problem described above. Take $D = \operatorname{im}(B \to C/I)$, which is a local artinian ring with residue field $K = B/\mathfrak{n}$. Next take $E = D \times_{C/I} C$. Then $E \to D$ is surjective and $E \subseteq C$. It remains to show that there is a lifting $B \to E$. If E was artinian, then we would be done by formal versality. But E need not be noetherian and we will instead construct an A-subalgebra $E_0 \subseteq E$ which is artinian and such that $E_0 \to E \to D$ is surjective with nilpotent kernel. Then $B \to D$ factors via A-homomorphisms $B \to E_0 \to E \to D$ by formal versality.

Let $k = A/\mathfrak{m}$ and first assume that $k \to K$ is a finitely generated extension. Since $E \to D \to K$ is surjective we may choose $t_1, \ldots, t_r \in E$ such that $k(t_1, \ldots, t_r) = K$. Further choose $u_1, \ldots, u_s \in E$ such that their images in D generate the maximal ideal. Let E_0 be the total quotient ring of the A-subalgebra of E generated by $t_1, \ldots, t_r, u_1, \ldots, u_s$. Then $E_0 \subseteq E$ is local artinian, $E_0 \to D$ is surjective, and by formal versality we have the required lift.

If instead $k \to K$ is separable, then there exists a Cohen A-algebra A' such that $A' \otimes_A k = K$. Recall that A' is a complete local noetherian ring and that $A \to A'$ is preadically formally smooth [EGA IV$_1$ 1964, Théorème 0.19.8.2]. Since $E \to D \to K$ is surjective with nilpotent kernel, we obtain a factorization $A \to A' \to E$ such that $A' \to E$ induces an isomorphism on residue fields. We can now take E_0 as the A'-subalgebra of E generated by $u_1, \ldots, u_s$. □

We now obtain the following well-known corollary (see [EGA IV$_4$ 1967, Proposition 17.14.2]).

Corollary 2.5. *Let $f : T \to X$ be a locally of finite type morphism of locally noetherian schemes. Let $t \in |T|$. The following are equivalent*:

(1) *f is smooth at t* [EGA IV$_4$ 1967, Définition 17.3.7, p. 62].

(2) *f is formally smooth at $t \in |T|$.*

(3) *f is formally versal at $t \in |T|$.*

Proof. Since f is locally of finite type, $\kappa(f(t)) \subseteq \kappa(t)$ is a finitely generated extension. By Proposition 2.4, it follows that conditions (2) and (3) are equivalent to $\mathcal{O}_{X,f(t)} \to \mathcal{O}_{T,t}$ being preadically formally smooth. By [EGA IV$_4$ 1967, Proposition 17.5.3], we have the claim. We can also argue as follows: the natural map $\operatorname{Exal}_X(T, \kappa(t)) \to \operatorname{Exal}_X(\operatorname{Spec} \mathcal{O}_{T,t}, \kappa(t))$ is an isomorphism. Indeed, the cotangent complex of the morphism $\operatorname{Spec} \mathcal{O}_{T,t} \to T$ vanishes. By Lemma 2.2, formal versality implies that $\operatorname{Exal}_X(\operatorname{Spec} \mathcal{O}_{T,t}, \kappa(t)) \cong 0$. By [Hall 2017, Lemma 5.4], the functor on quasicoherent $\mathcal{O}_T$-modules $\operatorname{Exal}_X(T, -)$ is coherent and limit preserving. By [Hall 2014, Corollary 7.7], there is thus an affine open neighborhood $j : U \subseteq T$ of t such that the functor $\operatorname{Exal}_X(T, j_*(-))$ vanishes. But $\operatorname{Exal}_X(T, j_*(-)) \simeq \operatorname{Exal}_X(U, -)$, so $U \to X$ is formally smooth [Hall 2017, Lemma 4.3(1)]. □

Corollary 2.6. *Let S be a locally noetherian scheme and let X be a limit preserving S-groupoid. Let T be an X-scheme that is locally of finite type over S and let $t \in |T|$ be a point such that*

(1) *T is formally smooth at $t \in |T|$ as an X-scheme and*

(2) *the morphism $T \to X$ is representable by algebraic spaces.*

*If W is an X-scheme, then the morphism $T \times_X W \to W$ is smooth in a neighborhood of every point over t. In particular, if $T \to X$ is formally smooth at every point of **finite type**, then $T \to X$ is formally smooth.*

Proof. By a standard limit argument, we can assume that $W \to S$ is of finite type. It is then enough to verify that $T \times_X W \to W$ is smooth at closed points in the fiber of t. Let $u\colon U \to T \times_X W$ be an étale and surjective morphism, where U is a scheme. Then $U \to W$ is formally smooth at closed points in the fiber of t. By Corollary 2.5, the composition $U \to W$ is smooth at every point over t, and we deduce the claim. The last statement follows from the fact that every closed point of $T \times_X W$ maps to a point of finite type of T. □

Combining Lemma 2.2 and Corollary 2.6 we obtain the following key result.

Corollary 2.7. *Let S be a locally noetherian scheme and let X be a limit preserving and $\mathbf{Art}^{\mathrm{fin}}$-semihomogeneous S-groupoid. If T is an X-scheme such that*

(1) *$T \to S$ is locally of finite type,*

(2) *$T \to X$ is formally versal at all points of finite type, and*

(3) *$T \to X$ is representable by algebraic spaces,*

then $T \to X$ is formally smooth.

Remark 2.8. To establish algebraicity of a functor or groupoid in the spirit of Artin's criteria, one must provide conditions for an algebraic family that is formally versal at all points of finite type to be formally smooth. In the present paper, this is Corollary 2.7, where we use $\mathbf{Art}^{\mathrm{fin}}$-semihomogeneity. This result was known to several experts. Artin [1969b, Lemma 5.4] also proved this result for fppf *sheaves* that are $\mathbf{Art}^{\mathrm{triv}}$-homogeneous. By Lemma 1.9, the fppf stack condition together with $\mathbf{Art}^{\mathrm{triv}}$-homogeneity imply $\mathbf{Art}^{\mathrm{fin}}$-homogeneity, so the results of our paper recover Artin's. As discussed in the Introduction, Artin's arguments for functors do not extend to groupoids.

For groupoids, the relationship between formal versality and smoothness is established in [Artin 1974, Proposition 4.2]. The relevant standing assumption is **rCl**-semihomogeneity. Assuming **rCl**-homogeneity makes no difference to our discussion below. We feel that it is worthwhile to digress into some of the technicalities that arise here. We wish to assure the reader that, as mentioned in the Introduction, if S is of finite type over $\operatorname{Spec}\mathbb{Z}$ or a perfect field, then the proof of the main result of [Artin 1974] is essentially correct, with only minor modifications to the arguments necessary.

Our interpretation of Artin's definition of formal smoothness [1974, p. 173] is that it coincides with ours given in Definition 2.1. In particular, in the notation of that work, to verify formal smoothness the residue fields of A are unconstrained. But the proof of [loc. cit., Proposition 4.2] relies on Theorem 3.3 in the same reference, which requires that the residue field of A is equal to the residue field of R (here both A and R are henselian local rings). If the residue field extension is separable, then it is possible

to conclude using [loc. cit., Proposition 4.3], which uses étale localization of obstruction theories (also see Proposition 2.9). We do not know how to complete the argument if the residue field extension is inseparable. The essential problem is the verification that formal versality is smooth-local.

It was suggested by a referee that Artin's definition of formal smoothness can be interpreted as follows. In the notation of [Artin 1974, p. 173], the morphism Spec $A \to$ Spec R should induce an isomorphism of residue fields at every point of finite type over S. With this definition of formal smoothness, Artin's proof of [loc. cit., Proposition 4.2] is correct. This definition of formal smoothness seems too limited to prove his main result [loc. cit., Corollary 5.2] without further assumptions, however. Indeed, it is essential in his Corollary 5.2 that formal smoothness is stable under base change. Artin omits the proof of this stability under base change and we were unable to prove it ourselves. Again, it is the presence of inseparable field extensions that complicates matters. Note that our definition of formal smoothness is obviously stable under base change.

2.1. ***Étale localization.*** We also obtain the following result showing that, under mild hypotheses, formal versality is stable under étale-localization. This improves [Artin 1974, Proposition 4.3], which requires the existence of an obstruction theory that is compatible with étale localization.

Proposition 2.9. *Let S be a scheme and let X be an* $\mathbf{Art}^{\mathrm{sep}}$*-semihomogeneous S-groupoid* (*see* Lemma 1.9). *Let T be an X-scheme. If $(U, u) \to (T, t)$ is a pointed étale morphism of S-schemes, then formal versality at $t \in |T|$ implies formal versality at $u \in |U|$.*

Proof. To see that formal versality at $t \in |T|$ implies formal versality at $u \in |U|$, it is enough to show that the lifting property holds for T and a square-zero extension of local artinian schemes $V \hookrightarrow V'$ such that $\kappa(v) = \kappa(u)$. This follows from an identical argument as in the proof of Lemma 2.2(2) $\Longrightarrow$ (1). □

Using Lemma 2.2, one can show that Proposition 2.9 admits a partial converse. Indeed, if $u \in |U|$ and $t \in |T|$ are closed, X is **rCl**-homogeneous, U and T are affine and noetherian, and $T \to X$ is representable by algebraic spaces, then formal versality at $u \in |U|$ implies formal versality at $t \in |T|$. This will not be used, however.

Remark 2.10. The conditions on obstruction theories in the criteria for algebraicity are used to prove that formal versality is an open condition. Proposition 2.9 proves that it is enough to find suitable obstruction theories étale-locally. This idea is present in [Artin 1974, 4.9–4.11]. We do not understand the given arguments, however, as they rely on [Artin 1974, Proposition 4.3], which requires the existence of a global obstruction theory. These are isolated remarks, however, having no bearing on the main results of the article.

2.2. ***Zariski localization.*** Next, we give a condition that ensures that if an X-scheme T is formally versal at all *closed* points, then it is formally versal at all points of *finite type*.

Condition 2.11. Let X be **Nil**-homogeneous and let T be an affine X-scheme. The extensions of X are *Zariski local at T* if for every open immersion $p\colon U \to T$ of affine X-schemes and every point $u \in |U|$

of finite type, the natural map:

$$\mathrm{Exal}_X(T, \kappa(u)) \to \mathrm{Exal}_X(U, \kappa(u))$$

is surjective. The extensions of X are *Zariski local* if they are Zariski local at every affine X-scheme that is locally of finite type over S.

Note that Lemma 1.8(5) implies that if an S-groupoid X is **Aff**-homogeneous, then its extensions are Zariski local. As the following lemma shows, it is also satisfied whenever S is Jacobson.

Lemma 2.12. *Let X be a* **Nil***-homogeneous Zariski S-stack and let $p\colon U \to T$ be an open immersion of affine X-schemes. If $u \in |U|$ is a point that is closed in T, then the natural map*:

$$\mathrm{Exal}_X(T, \kappa(u)) \to \mathrm{Exal}_X(U, \kappa(u))$$

is an isomorphism. In particular, if S is Jacobson, then extensions of X are Zariski local (Condition 2.11).

Proof. We construct an inverse by taking an X-extension $U \hookrightarrow U'$ of U by $\kappa(u)$ to the gluing of U' and $T \setminus \{u\}$ along $U' \setminus \{u\} \cong U \setminus \{u\}$. If S is Jacobson and $T \to S$ is locally of finite type, then T is Jacobson and every point of finite type $u \in |U|$ is closed in T so Condition 2.11 holds. □

We now extend the implication (3) $\implies$ (2) of Lemma 2.2 to points of finite type.

Proposition 2.13. *Fix a scheme S, an* **rCl***-homogeneous S-groupoid X and an affine X-scheme T, locally of finite type over S. Assume that extensions of X are Zariski local at T* (Condition 2.11). *If $t \in |T|$ is a point of* ***finite type*** *and $\mathrm{Exal}_X(T, \kappa(t)) = 0$, then the X-scheme T is formally versal at t.*

Proof. Finite type points are locally closed so there exists an open affine neighborhood $U \subseteq T$ of t such that $t \in |U|$ is closed. By Condition 2.11, $0 = \mathrm{Exal}_X(T, \kappa(t)) \twoheadrightarrow \mathrm{Exal}_X(U, \kappa(t))$, so the X-scheme U is formally versal at t by Lemma 2.2. It then follows, from the definition, that the X-scheme T also is formally versal at t. □

2.3. ***DVR-homogeneity.*** In this subsection, we will increase our homogeneity assumption instead of assuming that Exal commutes with localization.

Recall that a *geometric discrete valuation ring* is a discrete valuation ring D such that $\operatorname{Spec} D \to S$ is essentially of finite type and the residue field is of finite type over S [Artin 1969b, p. 38].

Notation 2.14. Let $\mathbf{DVR} \subseteq \mathbf{Aff}$ be the class of morphisms $(\operatorname{Spec} K \to \operatorname{Spec} D)$ such that D is a geometric discrete valuation ring with fraction field K.

Artin's condition [4a] of his Theorem 3.7 [1969b] implies **DVR**-semihomogeneity and Artin's conditions [5′](b) and [4′](a,b) of his Theorem 5.3 in the same work imply **DVR**-homogeneity. We conclude this section by showing that **DVR**-homogeneity implies that formal smoothness is stable under generizations. This is accomplished by the following lemma, which is a generalization of [Artin 1969b, Lemma 3.10] from functors to categories fibered in groupoids. To guarantee sufficiently many geometric discrete valuation rings, we assume that we are over an excellent base.

Lemma 2.15. *Let S be an excellent scheme and let X be a limit preserving* **DVR***-homogeneous S-groupoid. If T is an X-scheme such that*

(1) *$T \to S$ is locally of finite type,*

(2) *$T \to X$ is representable by algebraic spaces, and*

(3) *$T \to X$ is formally smooth at a point $t \in |T|$ of finite type,*

then $T \to X$ is formally smooth at every generization $t' \in |T|$ of t.

Proof. Consider a diagram of X-schemes

$$\begin{array}{ccc} Z_0 & \xrightarrow{g} & T \\ \downarrow & \nearrow & \downarrow \\ Z & \longrightarrow & X \end{array}$$

where $Z_0 \hookrightarrow Z$ is a closed immersion of local artinian schemes and the image $t' = g(z_0)$ of the closed point $z_0 \in |Z_0|$ is a generization of $t \in T$ and $\kappa(z_0)/\kappa(t')$ is finite. We have to prove that every such diagram admits a lifting as indicated by the dashed arrow.

As X is limit preserving, we can factor $Z \to X$ as $Z \to W \to X$ where W is an S-scheme of finite type. Let $h\colon T \times_X W \to T$ denote the first projection. The pull-back $T \times_X W \to W$ is smooth at every point of the fiber $h^{-1}(t)$ by Corollary 2.6. Let T_t denote the local scheme $\operatorname{Spec}(\mathcal{O}_{T,t})$. It is enough to prove that $T \times_X W \to W$ is smooth at every point of $h^{-1}(T_t)$.

Let $y \in |T \times_X W|$ be a point of $h^{-1}(T_t)$. It is enough to prove that $Y = \overline{\{y\}}$ contains a point at which $T \times_X W \to W$ is smooth. If $h(y) = t$, then we are done. If not, then by Chevalley's theorem, $h(Y)$ is indconstructible, hence contains a constructible neighborhood of $h(y)$. Thus, there is a point $t_1 \in h(Y) \cap T_t$ such that the closure $T_1 = \overline{\{t_1\}}$ in the local scheme T_t is of dimension 1. By Corollary 2.6, it is enough to show that $T \to X$ is formally smooth at t_1. Thus, consider a diagram

$$\begin{array}{ccc} \operatorname{Spec} K' & \xrightarrow{g} & T \\ \downarrow & \nearrow & \downarrow \\ \operatorname{Spec} K'' & \longrightarrow & X \end{array}$$

of X-schemes where $K'' \twoheadrightarrow K'$ is a surjection of local artinian rings such that $g(\eta) = t_1$ and $\kappa(\eta)/\kappa(t_1)$ is finite. Let $D \subseteq K = \kappa(\eta)$ be a geometric DVR dominating $\mathcal{O}_{T_1,t}$ (which exists since $\mathcal{O}_{T_1,t}$ is excellent). We may then, using **DVR**-homogeneity, extend the situation to a diagram

$$\begin{array}{ccccc} \operatorname{Spec} K' & \longrightarrow & \operatorname{Spec} D' & \longrightarrow & T \\ \downarrow & & \downarrow & \nearrow & \downarrow \\ \operatorname{Spec} K'' & \longrightarrow & \operatorname{Spec} D'' & \longrightarrow & X \end{array}$$

where $D' = D \times_K K'$ and $D'' = D \times_K K''$ so $D' \twoheadrightarrow D$ and $D'' \twoheadrightarrow D$ have nilpotent kernels. Now, by Corollary 2.6, the pullback $T \times_X \operatorname{Spec} D'' \to \operatorname{Spec} D''$ is smooth at the image of $\operatorname{Spec} D'$ so there is a lifting as indicated by the dashed arrow. Thus $T \to X$ is formally smooth at t_1 and hence also at t'. □

In Lemma 10.4 we will show that under mild hypotheses, **DVR**-homogeneity actually implies **Aff**-homogeneity and thus also Condition 2.11.

Remark 2.16. If we replace geometric DVRs with all DVRs in **DVR**-homogeneity, then it is enough that S is noetherian instead of excellent and t need not be of finite type.

3. Vanishing loci for additive functors

Let T be a scheme. In this section we will be interested in additive functors $F\colon \mathsf{QCoh}(T) \to \mathsf{Ab}$. It is readily seen that the collection of all such functors forms an abelian category, with all limits and colimits computed "pointwise". For example, given additive functors $F, G\colon \mathsf{QCoh}(T) \to \mathsf{Ab}$ as well as a natural transformation $\varphi\colon F \to G$, then $\ker\varphi\colon \mathsf{QCoh}(T) \to \mathsf{Ab}$ is the functor

$$(\ker\varphi)(M) = \ker(F(M) \xrightarrow{\varphi(M)} G(M)).$$

Next, we set $A = \Gamma(\mathcal{O}_T)$. Note that the natural action of A on the abelian category $\mathsf{QCoh}(T)$ induces for every $M \in \mathsf{QCoh}(T)$ an action of A on the abelian group $F(M)$. Thus we see that the functor F is canonically valued in the category $\mathsf{Mod}(A)$. It will be convenient to introduce the following notation: for a morphism between affine schemes $g\colon W \to T$ and a functor $F\colon \mathsf{QCoh}(T) \to \mathsf{Ab}$, define $F_W\colon \mathsf{QCoh}(W) \to \mathsf{Ab}$ to be the functor $F_W(N) = F(g_*N)$. If F is additive (resp. preserves direct limits), then the same is true of F_W. The *vanishing locus of* F is the following subset [Hall 2014, §7.2]:

$$\mathbb{V}(F) = \{t \in |T| : F_{\operatorname{Spec}(\mathcal{O}_{T,t})} \equiv 0\}.$$

The main result of this section, Theorem 3.3, which gives a criterion for the set $\mathbb{V}(F)$ to be Zariski open, is essentially due to H. Flenner. In [Flenner 1981, Lemma 4.1], for an S-groupoid X and an affine X-scheme V, locally of finite type over S, a specific result about the vanishing locus of the functor $M \mapsto \operatorname{Exal}_X(V, M)$ is proved. In that same work, a standing assumption is that the S-groupoid X is *semihomogeneous*, thus the functor $M \mapsto \operatorname{Exal}_X(T, M)$ is only set-valued, which complicates matters. Since we are assuming **Nil**-homogeneity of X, the functor $M \mapsto \operatorname{Exal}_X(T, M)$ takes values in abelian groups. As we will see, this simplifies matters considerably.

We now make the following trivial observation.

Lemma 3.1. *Let T be an affine scheme and let $F\colon \mathsf{QCoh}(T) \to \mathsf{Ab}$ be an additive functor. Then the subset $\mathbb{V}(F) \subseteq |T|$ is stable under generization.*

By Lemma 3.1, we thus see that the subset $\mathbb{V}(F) \subseteq |T|$ will be Zariski open if we can determine sufficient conditions on the functor F and the scheme T such that the subset $\mathbb{V}(F)$ is (ind)constructible. We make the following definitions:

Definition 3.2. Let $T = \operatorname{Spec} A$ be an affine scheme and let $F\colon \mathsf{QCoh}(T) \to \mathsf{Ab}$ be an additive functor.

- The functor F is *bounded* if the scheme T is noetherian and $F(M)$ is finitely generated for every finitely generated A-module M.
- The functor F is *weakly bounded* if the scheme T is noetherian and for every integral closed subscheme $T_0 \hookrightarrow T$, the $\Gamma(\mathcal{O}_{T_0})$-module $F(\mathcal{O}_{T_0})$ is finitely generated.
- The functor F is *GI* (resp. *GS*, resp. *GB*) if there exists a dense open subset $U \subseteq |T|$ such that for all points $u \in |U|$ of finite type, the natural map

$$F(\mathcal{O}_T) \otimes_A \kappa(u) \to F(\kappa(u))$$

  is injective (resp. surjective, resp. bijective).
- The functor F is *CI* (resp. *CS*, resp. *CB*) if for every integral closed subscheme $T_0 \hookrightarrow T$, the functor F_{T_0} is GI (resp. GS, resp. GB).

In the above definition, GI (resp. GS, resp. GB) is an acronym for *generically* injective (resp. surjective, resp. bijective). Similarly, CI (resp. CS, resp. CB) is an acronym for *constructibly* injective (resp. surjective, resp. bijective).

We can now state the main result of this section.

Theorem 3.3 (Flenner). *Let T be an affine noetherian scheme and let $F\colon \mathsf{QCoh}(T) \to \mathsf{Ab}$ be a half-exact, additive, and bounded functor that commutes with direct limits. If the functor F is CS, then the subset $\mathbb{V}(F) \subseteq |T|$ is Zariski open.*

Functors of the above type occur frequently in algebraic geometry.

Example 3.4. Let T be an affine noetherian scheme and let $Q \in \mathsf{D}^-_{\mathsf{Coh}}(T)$. Then, for all $i \in \mathbb{Z}$, the functors on quasicoherent $\mathcal{O}_T$-modules given by $M \mapsto \operatorname{Ext}^i_{\mathcal{O}_T}(Q, M)$ and $M \mapsto \operatorname{Tor}_i^{\mathcal{O}_T}(Q, M)$ are additive, bounded, half-exact, commute with direct limits, and CB.

Example 3.5. Let T be an affine noetherian scheme and let $p\colon X \to T$ be a morphism that is projective and flat. Then the functor $M \mapsto \Gamma(X, p^*M)$ is CB. Indeed, one interpretation of the Cohomology and Base Change Theorem asserts that the functor $M \mapsto \Gamma(X, p^*M)$ is of the form given in Example 3.4.

Example 3.6. Let T be an affine noetherian scheme. An additive functor $F\colon \mathsf{QCoh}(T) \to \mathsf{Ab}$, commuting with direct limits, is *coherent* [Auslander 1966] if there exists a homomorphism $M \to N$ of coherent $\mathcal{O}_T$-modules such that

$$F(-) = \operatorname{coker}(\operatorname{Hom}_{\mathcal{O}_T}(N, -) \longrightarrow \operatorname{Hom}_{\mathcal{O}_T}(M, -)).$$

It is easily seen that a coherent functor is CB and bounded. Indeed, boundedness is obvious and if $i\colon T_0 \hookrightarrow T$ is an integral closed subscheme, then $F|_{T_0} = \operatorname{coker}(\operatorname{Hom}_{\mathcal{O}_{T_0}}(i^*N, -) \to \operatorname{Hom}_{\mathcal{O}_{T_0}}(i^*M, -))$ and after passing to a dense open subscheme, we may assume that i^*N and i^*M are flat. Then $F|_{T_0}(-) =$

$\operatorname{coker}((i^*N)^\vee \to (i^*M)^\vee) \otimes_{\mathcal{O}_{T_0}} (-)$ commutes with all tensor products. It is well-known, and easily seen, that the functors of the previous two examples are coherent.

Conversely, let $F\colon \mathsf{QCoh}(T) \to \mathsf{Ab}$ be a half-exact bounded additive functor that commutes with direct limits and is CS. Then for every integral closed subscheme $T_0 \hookrightarrow T$, there is an affine open dense subscheme $U_0 \subseteq T_0$ such that such that $F|_{U_0}(-) = F(\mathcal{O}_{U_0}) \otimes -$, hence $F|_{U_0}$ is coherent. This follows from Theorem 3.3 and Proposition 3.9; see the proof of [Hall 2014, Corollary 7.8]. In particular, for half-exact bounded additive functors that commute with direct limits, CS implies CB.

The main ingredient in the proof of Theorem 3.3 is a remarkable Nakayama lemma for half-exact functors, due to A. Ogus and G. Bergman [1972, Theorem 2.1]. We state the following amplification, which follows from the mild strengthening given in [Hall 2014, Corollary 7.5] and Lemma 3.1.

Theorem 3.7 (Ogus–Bergman). *Let T be an affine noetherian scheme and let $F\colon \mathsf{QCoh}(T) \to \mathsf{Ab}$ be a half-exact, additive, and bounded functor that commutes with direct limits. Then*

$$\mathbb{V}(F) = \{t \in |T| : F(\kappa(t)) = 0\}.$$

In particular, if $F(\kappa(t)) = 0$ for all closed points $t \in |T|$, then $F \equiv 0$.

Remark 3.8. Let F be as in Theorem 3.7 and let $I \subseteq A$ be an ideal. Then Flenner proves that the natural map $F(M) \otimes_A \hat{A}_{/I} \to \varprojlim_n F(M/I^nM)$ is injective for every finitely generated A-module M. In fact, this is the special case $X = Y = \operatorname{Spec} A$ of [Flenner 1981, Korollar 6.3]. The Ogus–Bergman–Nakayama lemma is an immediate consequence of the injectivity of this map.

Before we address vanishing loci of functors, the following simple application of Lazard's theorem [1964], which appeared in [Hall 2014, Proposition 7.2], will be a convenient tool to have at our disposal.

Proposition 3.9. *Let $T = \operatorname{Spec} A$ be an **affine** scheme and let $F\colon \mathsf{QCoh}(T) \to \mathsf{Ab}$ be an additive functor that commutes with direct limits. Let M and L be A-modules. If L is **flat**, then the natural map*

$$F(M) \otimes_A L \to F(M \otimes_A L)$$

is an isomorphism. In particular, for every A-algebra B and every flat B-module L, the natural map

$$F(B) \otimes_B L \to F(L)$$

is an isomorphism.

We may now prove Flenner's theorem.

Proof of Theorem 3.3. The subset $\mathbb{V}(F) \subseteq |T|$ is open if and only if it is closed under generization and its intersection with any irreducible closed subset $T_0 \subseteq |T|$ contains a nonempty open subset of T_0 or is empty [EGA IV$_1$ 1964, Théorème 1.10.1]. By Lemma 3.1, we have witnessed the stability under generization. Thus it remains to address the latter claim.

Let $T_0 \hookrightarrow T$ be an integral closed subscheme. If $|T_0| \cap \mathbb{V}(F) \neq \varnothing$, then the generic point $\eta \in |T_0|$ belongs to $\mathbb{V}(F)$ (Lemma 3.1), thus $F(\kappa(\eta)) = 0$. Since by assumption the functor F is CS, there exists

a dense open subset $U_0 \subseteq |T_0|$ such that the map $F_{T_0}(\mathcal{O}_{T_0}) \otimes_{\Gamma(\mathcal{O}_{T_0})} \kappa(u) \to F(\kappa(u))$ is surjective for all $u \in U_0$ of finite type.

As $\kappa(\eta)$ is a quasicoherent and flat $\mathcal{O}_{T_0}$-module, the natural map $F_{T_0}(\mathcal{O}_{T_0}) \otimes_{\Gamma(\mathcal{O}_{T_0})} \kappa(\eta) \to F(\kappa(\eta))$ is an isomorphism by Proposition 3.9. But $\eta \in \mathbb{V}(F)$, thus the finitely generated $\Gamma(\mathcal{O}_{T_0})$-module $F_{T_0}(\mathcal{O}_{T_0})$ is torsion. Hence there is a dense open subset $U_0 \subseteq |T_0|$ with the property that if $u \in U_0$ is of finite type, then $F(\kappa(u)) = 0$. Using Theorem 3.7 we infer that $U_0 \subseteq \mathbb{V}(F) \cap |T_0|$. □

We record for future reference a useful lemma.

Lemma 3.10. *Let $T = \operatorname{Spec} A$ be an affine noetherian scheme and let $F\colon \mathsf{QCoh}(T) \to \mathsf{Ab}$ be an additive functor.*

(1) *If the functor F is **half-exact**, then F is bounded if and only if F is weakly bounded.*

(2) *If the functor F is (weakly) bounded, then every additive subquotient functor of F is (weakly) bounded.*

(3) *If F is GS (resp. CS), then so is every additive quotient functor of F.*

(4) *If F is weakly bounded and CI, then so is every additive subfunctor of F.*

(5) *Consider an exact sequence of additive functors $\mathsf{QCoh}(T) \to \mathsf{Ab}$:*

$$H_1 \longrightarrow H_2 \longrightarrow H_3 \longrightarrow H_4.$$

(a) *If H_1 and H_3 are CS and H_4 is CI and weakly bounded, then H_2 is CS.*

(b) *If H_1 is CS, H_2 and H_4 are CI, and H_4 is weakly bounded, then H_3 is CI.*

If T is reduced, then (4), (5a), and (5b) hold with GI and GS instead of CI and CS.

Proof. For claim (1), note that every coherent $\mathcal{O}_T$-module M admits a finite filtration whose successive quotients are of the form $i_*\mathcal{O}_{T_0}$, where $i\colon T_0 \hookrightarrow T$ is an integral closed subscheme. Induction on the length of the filtration, combined with the half-exactness of the functor F, proves the claim. Claims (2) and (3) are trivial. For (4), it is sufficient to prove the claim about GI and we can assume that T is a disjoint union of integral schemes. Fix an additive subfunctor $K \subseteq F$, then there is an exact sequence of additive functors: $0 \to K \to F \to H \to 0$. By (2) we see that H is weakly bounded and so $H(\mathcal{O}_T)$ is a finitely generated A-module. As A is reduced, generic flatness implies that there is a dense open subset $U \subseteq |T|$ such that $H(\mathcal{O}_T)_u$ is a flat A-module $\forall u \in U$. Thus, for all $u \in U$ the sequence

$$0 \longrightarrow K(\mathcal{O}_T) \otimes_A \kappa(u) \longrightarrow F(\mathcal{O}_T) \otimes_A \kappa(u) \longrightarrow H(\mathcal{O}_T) \otimes_A \kappa(u) \longrightarrow 0$$

is exact. Since F is GI, we may further assume that the map $F(\mathcal{O}_T) \otimes_A \kappa(u) \to F(\kappa(u))$ is injective for all points $u \in U$ of finite type after shrinking U. We then conclude that K is GI from the commutative diagram

$$\begin{array}{ccc} K(\mathcal{O}_T) \otimes_A \kappa(u) & \hookrightarrow & F(\mathcal{O}_T) \otimes_A \kappa(u) \\ \downarrow & & \downarrow \\ K(\kappa(u)) & \hookrightarrow & F(\kappa(u)) \end{array}$$

Claims (5a) and (5b) follow from a similar argument and the 4-Lemmas. □

We conclude this section with a criterion for a functor to be GI (and consequently a criterion for a functor to be CI). This will be of use when we express Artin's criteria for algebraicity without obstruction theories in Section 8.

Proposition 3.11. *Let $T = \operatorname{Spec} A$ be an affine and integral (i.e., reduced and irreducible) noetherian scheme with function field K. Let $F\colon \mathsf{QCoh}(T) \to \mathsf{Ab}$ be an additive functor that commutes with direct limits. If $F(\mathcal{O}_T)$ is a finitely generated A-module, then F is GI if and only if the following condition is satisfied*:

(†) *for every $f \in A$, every free A_f-module M of finite rank, and $\omega \in F(M)$ such that for all nonzero A-module maps $\epsilon\colon M \to K$ we have $\epsilon_*\omega \neq 0$ in $F(K)$, there exists a dense open subset $V_\omega \subseteq D(f) \subseteq |T|$ such that for every nonzero A-module map $\gamma\colon M \to \kappa(v)$, where $v \in V_\omega$ is of finite type, we have $\gamma_*\omega \neq 0$ in $F(\kappa(v))$.*

Proof. Let M be a free A_f-module of finite rank and let $M^\vee = \operatorname{Hom}_{A_f}(M, A_f)$. Then the canonical homomorphism $F(A)_f \otimes_{A_f} M \to F(M)$ is an isomorphism (Proposition 3.9) so there is a one-to-one correspondence between elements $\omega \in F(M)$ and homomorphisms $\bar\omega\colon M^\vee \to F(A)_f$. Moreover, $\bar\omega$ is injective if and only if $\bar\omega \otimes_A K\colon M^\vee \otimes_A K \to F(A) \otimes_A K = F(K)$ is injective and this happens exactly when $\epsilon_*\omega \neq 0$ in $F(K)$ for every nonzero map $\epsilon\colon M \to K$.

Let $t \in |T|$ and let $\delta_t\colon F(A) \otimes_A \kappa(t) \to F(\kappa(t))$ denote the natural map. Then condition (†) can be reformulated as: for every free A_f-module M of finite rank and every injective homomorphism $\bar\omega\colon M^\vee \to F(A)_f$, there exists a dense open subset $V_\omega \subseteq D(f)$ such that $\delta_t \circ (\bar\omega \otimes_A \kappa(t))$ is injective for all points $t \in V_\omega$ of finite type.

To show that (†) implies that F is GI, choose $f \in A \setminus 0$ such that $F(A)_f$ is free, let $M = F(A)_f^\vee$ and let $\omega \in F(M)$ correspond to the inverse of the canonical isomorphism $F(A)_f \to M^\vee$. If (†) holds, then there exists an open subset V_ω such that δ_t is injective for all $t \in V_\omega$, i.e., F is GI.

Conversely, if F is GI, then there is an open subset V such that δ_t is injective for all $t \in V$ of finite type. Given a finite free A_f-module M and $\omega \in F(M)$, we let $V_\omega = V \cap W$ where $W \subseteq D(f)$ is an open dense subset over which the cokernel of $\bar\omega$ is flat. If $\bar\omega$ is injective, it then follows that $\delta_t \circ (\bar\omega \otimes_A \kappa(t))$ is injective for all $t \in V_\omega$ of finite type, that is, condition (†) holds. □

4. Openness of formal versality

As the title suggests, we now address the openness of the formally versal locus. Let S be a scheme. We isolate the following conditions for an S-groupoid X.

Condition 4.1. Let T be an affine X-scheme. The extensions of X are *bounded at T* if X is **Nil**-homogeneous at T and the functor $M \mapsto \operatorname{Exal}_X(T, M)$ is bounded. The extensions of X are *bounded* if X has bounded extensions at every affine X-scheme T, locally of finite type over S.

Condition 4.2. Let T be an affine X-scheme. The extensions of X are *constructible at* T if X is **Nil**-homogeneous at T and the functor $M \mapsto \mathrm{Exal}_X(T, M)$ is CS. The extensions of X are *constructible* if X has constructible extensions at every affine X-scheme T, locally of finite type over S.

That these conditions are plausible is implied by the following lemma:

Lemma 4.3. *Let S be a locally noetherian scheme, let X be an algebraic S-stack, and let T be an affine X-scheme. If both X and T are locally of finite type over S, then the functors $M \mapsto \mathrm{Der}_X(T, M)$ and $M \mapsto \mathrm{Exal}_X(T, M)$ are bounded and CB.*

Proof. By [Olsson 2006, Theorem 1.1] there is a complex $L_{T/X} \in \mathsf{D}^-_{\mathsf{Coh}}(T)$ such that for all quasicoherent $\mathcal{O}_T$-modules M, there are natural isomorphisms $\mathrm{Der}_X(T, M) \cong \mathrm{Ext}^0_{\mathcal{O}_T}(L_{T/X}, M)$ and $\mathrm{Exal}_X(T, M) \cong \mathrm{Ext}^1_{\mathcal{O}_T}(L_{T/X}, M)$. The result now follows from a consideration of Example 3.4. □

In their current form, Conditions 4.1 and 4.2 are difficult to verify. In Section 6, this will be rectified. Nonetheless, we can now prove the following.

Theorem 4.4. *Let S be a locally noetherian scheme, let X be an S-groupoid and let T be an affine X-scheme, locally of finite type over S. Assume, in addition, that*

(1) *X is limit preserving,*

(2) *X is* **rCl***-homogeneous,*

(3) *X has bounded extensions at T* (Condition 4.1),

(4) *X has constructible extensions at T* (Condition 4.2), *and*

(5) *X has Zariski local extensions at T* (Condition 2.11).

Let $t \in |T|$ be a ***closed*** *point. If T is formally versal at $t \in |T|$, then T is formally versal at every point of finite type in a Zariski open neighborhood of t. In particular, if X is also $\mathbf{Art}^{\mathrm{fin}}$-homogeneous and $T \to X$ is representable by algebraic spaces, then T is formally smooth in a Zariski open neighborhood of t.*

Proof. By Condition 4.1 and Lemma 1.8, the functor $M \mapsto \mathrm{Exal}_X(T, M)$ is bounded, half-exact, and preserves direct limits. Condition 4.2 now implies that the functor $M \mapsto \mathrm{Exal}_X(T, M)$ satisfies the criteria of Theorem 3.3. Thus, $\mathbb{V}(\mathrm{Exal}_X(T, -)) \subseteq |T|$ is a Zariski open subset. By Lemma 2.2(2) $\Longrightarrow$ (3) and Theorem 3.7, we have that $t \in \mathbb{V}(\mathrm{Exal}_X(T, -))$. So, there exists an open neighborhood $t \in U \subseteq |T|$ with $\mathrm{Exal}_X(T, \kappa(u)) = 0$ for all $u \in U$. By Proposition 2.13, every point $u \in U$ of finite type is formally versal. The last assertion follows from Corollary 2.7. □

5. Automorphisms, deformations, and obstructions

In this section, we introduce a deformation-theoretic framework that makes it possible to verify Conditions 2.11, 4.1, and 4.2. To do this, we recall the formulation of deformations and obstructions given in [Hall 2017, §6].

Let S be a scheme and let $\Phi\colon Y \to Z$ be a 1-morphism of S-groupoids. Define the category $\mathbf{Def}_\Phi$ to have objects the pairs $(i\colon T \hookrightarrow T', r\colon T' \to T)$, where i is a Y-extension and r is a Z-retraction of i, with the obvious morphisms. Graphically, it is the category of completions of the following diagram:

$$\begin{array}{ccc} T & \longrightarrow & Y \\ \downarrow & \overset{\eta}{\nearrow} & \downarrow{\scriptstyle \Phi} \\ T[J] & \longrightarrow & Z \end{array}$$

Forgetting the retraction, there is a natural functor $\mathbf{Def}_\Phi \to \mathbf{Exal}_Y$. If T is a Y-scheme, then we denote the fiber of this functor over $\mathbf{Exal}_Y(T) \subseteq \mathbf{Exal}_Y$ by $\mathbf{Def}_\Phi(T)$. It follows that there is an induced functor $\mathbf{Def}_\Phi(T) \to \mathsf{QCoh}(T)^\circ$, whose fiber over a quasicoherent $\mathcal{O}_T$-module I we denote by $\mathbf{Def}_\Phi(T, I)$. Note that the category $\mathbf{Def}_\Phi(T, I)$ is naturally pointed by the trivial Y-extension $i_{T,J}$ of T by J. Denote the set of isomorphism classes of $\mathbf{Def}_\Phi(T, J)$ by $\mathrm{Def}_\Phi(T, J)$ and let $\mathrm{Aut}_\Phi(T, J)$ denote the set $\mathrm{Aut}_{\mathbf{Def}_\Phi(T,J)}(i_{T,J})$.

If Y and Z are **Nil**-homogeneous at T, then the groupoid $\mathbf{Def}_\Phi(T, J)$ is a Picard category [Hall 2017, Proposition 6.5]. Thus we obtain $\Gamma(T, \mathcal{O}_T)$-linear functors

$$\mathrm{Def}_\Phi(T, -)\colon \mathsf{QCoh}(T) \to \mathsf{Ab}, \quad J \mapsto \mathrm{Def}_\Phi(T, J); \quad \text{and}$$
$$\mathrm{Aut}_\Phi(T, -)\colon \mathsf{QCoh}(T) \to \mathsf{Ab}, \quad J \mapsto \mathrm{Aut}_{\mathbf{Def}_\Phi(T,J)}(i_{T,J}).$$

The lemma that follows is an easy consequence of [Hall 2017, Lemma 6.4].

Lemma 5.1. *Let S be a scheme and let $\Phi\colon Y \to Z$ be a 1-morphism S-groupoids. Let $i\colon W \hookrightarrow T$ be a closed immersion of Y-schemes and let N be a quasicoherent $\mathcal{O}_W$-module. If Y and Z are* **Cl***-homogeneous at W, then the natural maps*

$$\mathrm{Aut}_\Phi(T, i_*N) \to \mathrm{Aut}_\Phi(W, N) \quad \textit{and} \quad \mathrm{Def}_\Phi(T, i_*N) \to \mathrm{Def}_\Phi(W, N)$$

are bijective.

We recall the exact sequence of [Hall 2017, Proposition 6.7], which is our fundamental computational tool.

Proposition 5.2. *Let S be a scheme and let $\Phi\colon Y \to Z$ be a 1-morphism of S-groupoids. Let T be a Y-scheme and let J be a quasicoherent $\mathcal{O}_T$-module. If Y and Z are* **Nil***-homogeneous at T, then there is a natural 6-term exact sequence of abelian groups*

$$0 \longrightarrow \mathrm{Aut}_\Phi(T, J) \longrightarrow \mathrm{Der}_Y(T, J) \longrightarrow \mathrm{Der}_Z(T, J) \longrightarrow \mathrm{Def}_\Phi(T, J) \longrightarrow \mathrm{Exal}_Y(T, J) \longrightarrow \mathrm{Exal}_Z(T, J)$$

If Y and Z are **Nil**-homogeneous at T and J is a quasicoherent $\mathcal{O}_T$-module, then we let

$$\mathrm{Obs}_\Phi(T, J) = \mathrm{coker}(\mathrm{Exal}_Y(T, J) \to \mathrm{Exal}_Z(T, J)).$$

This defines a $\Gamma(T, \mathcal{O}_T)$-linear functor

$$\mathrm{Obs}_\Phi(T, -)\colon \mathsf{QCoh}(T) \to \mathsf{Ab}, \quad J \mapsto \mathrm{Obs}_\Phi(T, J),$$

the *minimal obstruction theory* of Φ at T (see Section 7). If Y and Z are **rNil**-homogeneous at T, then $\mathrm{Aut}_\Phi(T, -)$ and $\mathrm{Def}_\Phi(T, -)$ are half-exact [Hall 2017, Corollary 6.6]. There is no reason to expect that $\mathrm{Obs}_\Phi(T, -)$ is half-exact, however. We have the following analogues of Lemmas 1.8(2) and 5.1 for obstructions.

Lemma 5.3. *Let S be a scheme and let P be a class of morphisms of S-schemes. Let $\Phi\colon Y \to Z$ be a 1-morphism of S-groupoids. Let $p\colon V \to T$ be an affine morphism of Y-schemes that is P. If Y and Z are P-homogeneous at V and* **Nil***-homogeneous at T, then there is a natural map $p_\#\colon \mathrm{Obs}_\Phi(V, N) \to \mathrm{Obs}_\Phi(T, p_*N)$, which is injective and functorial in N.*

Proof. The existence of $p_\#$ follows immediately from Lemma 1.8(2). That $p_\#$ is injective is obvious. □

Lemma 5.4. *Let S be a scheme, and let $\Phi\colon Y \to Z$ be a 1-morphism of* **Cl***-homogeneous S-groupoids. Let $i\colon W \hookrightarrow T$ be a closed immersion of affine noetherian Y-schemes and let N be a quasicoherent $\mathcal{O}_W$-module. If $\mathrm{Obs}_\Phi(T, i_*N)$ is a finitely generated $\Gamma(T, \mathcal{O}_T)$-module, then there exists an infinitesimal neighborhood $i_n\colon W_n \to T$ of W in T, i.e., a factorization of i as $W \xrightarrow{j} W_n \xrightarrow{i_n} T$, where j is a locally nilpotent closed immersion, such that*

$$(i_n)_\#\colon \mathrm{Obs}_\Phi(W_n, j_*N) \to \mathrm{Obs}_\Phi(T, i_*N)$$

is an isomorphism.

Proof. Given an obstruction $\omega \in \mathrm{Obs}_\Phi(T, i_*N)$, we can realize it as a Z-extension $k\colon T \hookrightarrow T'$ of T by i_*N. The ideal sheaf $k_*i_*N \subseteq \mathcal{O}_{T'}$ is then annihilated by the ideal sheaf I defining the closed immersion $k \circ i\colon W \hookrightarrow T'$. Thus, by the Artin–Rees lemma, we have that $(k_*i_*N) \cap I^n = 0$ for some n. Let W_1' and W_1 be the closed subschemes of T' defined by I^n and $I^n + k_*i_*N$. Then the morphisms in the diagram

$$\begin{array}{ccccc} W & \xhookrightarrow{j_1} & W_1 & \xhookrightarrow{i_1} & T \\ & & \downarrow & & \downarrow \\ & & W_1' & \hookrightarrow & T' \end{array}$$

are closed immersions and the square is cartesian and cocartesian in the category of Z-schemes (because Z is **Cl**-homogeneous at W_1). If we let $\omega_1 = [W_1 \hookrightarrow W_1'] \in \mathrm{Obs}_\Phi(W_1, (j_1)_*N)$ denote the obstruction to lifting W_1' to a Y-scheme; then $\omega = (i_1)_\#(\omega_1)$.

We have thus shown that every element $\omega \in \mathrm{Obs}_\Phi(T, i_*N)$ is in the image of $\mathrm{Obs}_\Phi(W_l, (j_l)_*N)$ for some infinitesimal neighborhood $j_l\colon W \hookrightarrow W_l$, depending on ω. Since $\mathrm{Obs}_\Phi(T, i_*N)$ is a finitely generated $\Gamma(T, \mathcal{O}_T)$-module and T is affine and noetherian, it follows that there exists an infinitesimal neighborhood $j\colon W \hookrightarrow W_n$ such that $\mathrm{Obs}_\Phi(W_n, j_*N) \to \mathrm{Obs}_\Phi(T, i_*N)$ is an isomorphism. □

6. Relative conditions

Let S be a locally noetherian scheme. In this section, we introduce a number of conditions for a 1-morphism of S-groupoids $\Phi\colon Y \to Z$. These are the relative versions of the conditions that appear in (5a), (5b), (5c), (6b), and (6c) of the Main Theorem. For any of the conditions given in this section, an S-groupoid X is said to have that condition if the structure 1-morphism $X \to \mathsf{Sch}/S$ has the condition. These conditions are stated "relatively" for two reasons. The first reason is to make it clear that this paper subsumes the results of [Starr 2006] on the stability of Artin's criteria under composition. This follows immediately from the exact sequence of [Hall 2017, Proposition 6.13] and Lemma 3.10. Secondly, and of most importance, is that the relative formulation permits a process of bootstrapping the diagonal. This is an important and subtle point of this paper, which we will discuss in more detail when we prove the Main Theorem in Section 10.

Condition 6.1. Let T be an affine Y-scheme. Assume that Y and Z are **Nil**-homogeneous at every closed subscheme of T. Automorphisms (resp. deformations, resp. obstructions) of Φ are *bounded at* T if for every integral closed subscheme $i\colon T_0 \hookrightarrow T$, condition (i) (resp. (ii), resp. (iii)) below holds:

(i) $\mathrm{Aut}_\Phi(T_0, \mathcal{O}_{T_0})$ is a finitely generated $\Gamma(\mathcal{O}_{T_0})$-module;

(ii) $\mathrm{Def}_\Phi(T_0, \mathcal{O}_{T_0})$ is a finitely generated $\Gamma(\mathcal{O}_{T_0})$-module;

(iii) $\mathrm{Obs}_\Phi(T, i_*\mathcal{O}_{T_0})$ is a finitely generated $\Gamma(\mathcal{O}_{T_0})$-module.

Automorphisms (resp. deformations, resp. obstructions) of Φ are *bounded* if they are bounded at every affine Y-scheme T, locally of finite type over S.

Morphisms of S-groupoids typically have bounded obstructions (Condition 6.1(iii)). For example, if Y is **Nil**-homogeneous and Z is algebraic, then Z has bounded extensions (Condition 4.1) and Φ has bounded obstructions.

Lemma 6.2. *Let S be a locally noetherian scheme and let $\Phi\colon Y \to Z$ be a* 1*-morphism of* **rCl***-homogeneous S-groupoids with bounded deformations* (Condition 6.1(ii)) *at an affine Y-scheme T, locally of finite type over S. If Z has bounded extensions at T* (Condition 4.1)*, then so does Y.*

Proof. By Lemma 1.8(3) the functor $M \mapsto \mathrm{Exal}_Y(T, M)$ is half-exact. Thus, by Lemma 3.10(1), it is sufficient to prove that for every integral closed subscheme $i\colon T_0 \hookrightarrow T$, the $\Gamma(\mathcal{O}_{T_0})$-module $\mathrm{Exal}_Y(T, i_*\mathcal{O}_{T_0})$ is finitely generated. Now, by Proposition 5.2, there is an exact sequence

$$\mathrm{Def}_\Phi(T, i_*\mathcal{O}_{T_0}) \longrightarrow \mathrm{Exal}_Y(T, i_*\mathcal{O}_{T_0}) \longrightarrow \mathrm{Exal}_Z(T, i_*\mathcal{O}_{T_0}).$$

By Condition 4.1, the $\Gamma(\mathcal{O}_{T_0})$-module $\mathrm{Exal}_Z(T, i_*\mathcal{O}_{T_0})$ is finitely generated. By Lemma 5.1,

$$\mathrm{Def}_\Phi(T, i_*\mathcal{O}_{T_0}) \cong \mathrm{Def}_\Phi(T_0, \mathcal{O}_{T_0}),$$

which is also a finitely generated $\Gamma(\mathcal{O}_{T_0})$-module by Condition 6.1(ii). The result now follows from the exact sequence above. □

Similarly, to enable the verification that an S-groupoid has constructible extensions (Condition 4.2), we introduce the following conditions.

Condition 6.3. Let T be an affine Y-scheme. Assume that Y and Z are **Nil**-homogeneous at every closed subscheme of T. Automorphisms (resp. deformations, resp. obstructions) of Φ are *constructible at* T if for every closed subscheme $T_1 \subseteq T$, such that T_1 is irreducible and $i : T_0 \hookrightarrow T_1$ denotes the reduction, condition (i) (resp. (ii), resp. (iii)) below holds:

(i) $\mathrm{Aut}_\Phi(T_0, -) : \mathsf{QCoh}(T_0) \to \mathsf{Ab}$ is GB;

(ii) $\mathrm{Def}_\Phi(T_0, -) : \mathsf{QCoh}(T_0) \to \mathsf{Ab}$ is GB;

(iii) $\mathrm{Obs}_\Phi(T_1, i_*-) : \mathsf{QCoh}(T_0) \to \mathsf{Ab}$ is GI.

Automorphisms (resp. deformations, resp. obstructions) of Φ are *constructible* if they are constructible at every affine Y-scheme T, locally of finite type over S.

We now proceed to Zariski local extensions (Condition 2.11). Note that the following condition trivially holds when S is Jacobson. Indeed, in that case, $U_1 = T_1 = \{\eta\}$.

Condition 6.4. Let T be an affine Y-scheme. Assume that Y and Z are **Nil**-homogeneous at every closed subscheme of T. Automorphisms (resp. deformations, resp. obstructions) of Φ are *Zariski local at* T if for every closed subscheme $T_1 \subseteq T$ and nonempty open subscheme $U_1 \subseteq T_1$, such that T_1 is irreducible and the generic point $\eta \in |T_1|$ is of finite type over S, and $U_0 \subseteq T_0$ denotes the reductions, condition (i) (resp. (ii), resp. (iii)) below holds:

(i) the natural map $\mathrm{Aut}_\Phi(T_0, \kappa(\eta)) \to \mathrm{Aut}_\Phi(U_0, \kappa(\eta))$ is bijective;

(ii) the natural map $\mathrm{Def}_\Phi(T_0, \kappa(\eta)) \to \mathrm{Def}_\Phi(U_0, \kappa(\eta))$ is bijective;

(iii) the natural map $\mathrm{Obs}_\Phi(T_1, \kappa(\eta)) \to \mathrm{Obs}_\Phi(U_1, \kappa(\eta))$ is injective.

Automorphisms (resp. deformations, resp. obstructions) of Φ are *Zariski local* if they are Zariski local at every affine Y-scheme T, locally of finite type over S.

The following proposition is one of the major results of the article.

Proposition 6.5. *Let S be a locally noetherian scheme. Let $\Phi : Y \to Z$ be a 1-morphism of* **Cl**-*homogeneous S-groupoids with bounded obstructions at an affine Y-scheme T, locally of finite type over S* (Condition 6.1(iii)).

(1) *Assume, in addition, that Φ has constructible deformations and obstructions at T* (Conditions 6.3(ii)–(iii)). *If Z has constructible extensions at T* (Condition 4.2), *then so does Y.*

(2) *Assume, in addition, that Φ has Zariski local deformations and obstructions at T* (Conditions 6.4(ii)–(iii)). *If Z has Zariski local extensions at T* (Condition 2.11), *then so does Y.*

Proof. We prove (1). By Proposition 5.2 there is an exact sequence of additive functors $\mathsf{QCoh}(T) \to \mathsf{Ab}$

$$\mathrm{Def}_\Phi(T, -) \longrightarrow \mathrm{Exal}_Y(T, -) \longrightarrow \mathrm{Exal}_Z(T, -) \longrightarrow \mathrm{Obs}_\Phi(T, -) \longrightarrow 0.$$

Let $i\colon T_0 \hookrightarrow T$ be an integral closed subscheme. By Lemma 5.1 we have $\mathrm{Def}_\Phi(T_0, -) = \mathrm{Def}_\Phi(T, i_*(-))$. Condition 6.3(ii) gives that $\mathrm{Def}_\Phi(T_0, -)$ is GS, so the functor $\mathrm{Def}_\Phi(T, -)$ is CS. Condition 4.2 says that $\mathrm{Exal}_Z(T, -)$ is CS. The remaining two conditions together with Lemma 5.4 imply that $\mathrm{Obs}_\Phi(T, -)$ is CI and weakly bounded. In fact, for every integral closed subscheme $i\colon T_0 \hookrightarrow T$, there is an infinitesimal neighborhood $j\colon T_0 \hookrightarrow T_1$ such that $\mathrm{Obs}_\Phi(T_1, j_*\mathcal{O}_{T_0}) \cong \mathrm{Obs}_\Phi(T, i_*\mathcal{O}_{T_0})$ and $\mathrm{Obs}_\Phi(T_1, \kappa(t)) \hookrightarrow \mathrm{Obs}_\Phi(T, \kappa(t))$ is injective for all points t of finite type in a dense open subset of T_0. It now follows from Lemma 3.10(5a) that the functor $\mathrm{Exal}_Y(T, -)$ is CS.

The proof of (2) is similar: let $u \in U \subseteq T$ be as in Condition 2.11, use the exact sequence above, take $T_0 = \overline{\{u\}}$, and apply Lemmas 5.1 and 5.4 as before. □

7. Obstruction theories

Throughout this section, we let S be a locally noetherian scheme and let $\Phi\colon Y \to Z$ be a 1-morphism of **Nil**-homogeneous S-groupoids. In this section, we will expand the conditions on obstructions given in the previous sections to obtain more readily verifiable conditions. We begin with recalling the definition of an n-step relative obstruction theory given in [Hall 2017, Definition 6.8].

An *n-step relative obstruction theory for* Φ, denoted $\{\mathrm{o}^l(-,-), \mathrm{O}^l(-,-)\}_{l=1}^n$, is for each Y-scheme T, a sequence of additive functors (the obstruction spaces)

$$\mathrm{O}^l(T, -)\colon \mathsf{QCoh}(T) \to \mathsf{Ab}, \quad J \mapsto \mathrm{O}^l(T, J), \quad l = 1, \dots, n,$$

as well as natural transformations of functors (the obstruction maps)

$$\begin{aligned} \mathrm{o}^1(T, -)\colon\ & \mathrm{Exal}_Z(T, -) \Rightarrow \mathrm{O}^1(T, -), \\ \mathrm{o}^l(T, -)\colon\ & \ker \mathrm{o}^{l-1}(T, -) \Rightarrow \mathrm{O}^l(T, -) \quad \text{for } l = 2, \dots, n, \end{aligned}$$

such that the natural transformation of functors

$$\mathrm{Exal}_Y(T, -) \Rightarrow \mathrm{Exal}_Z(T, -)$$

has image $\ker \mathrm{o}^n(T, -)$. Furthermore, we say that the obstruction theory is

- *(weakly) bounded*, if for every affine Y-scheme T, locally of finite type over S, the obstruction spaces $M \mapsto \mathrm{O}^l(T, M)$ are (weakly) bounded functors;
- *Zariski- (resp. étale-) functorial* if for every open immersion (resp. étale morphism) of affine Y-schemes $g\colon V \to T$, and $l = 1, \dots, n$, there is a natural transformation of functors

$$C_g^l\colon \mathrm{O}^l(T, g_*(-)) \Rightarrow \mathrm{O}^l(V, -),$$

which for every quasicoherent $\mathcal{O}_V$-module N, make the following diagrams commute:

$$\begin{array}{ccc} \mathrm{Exal}_X(T, g_*N) & \longrightarrow & \mathrm{O}^1(T, g_*N) \\ \downarrow & & \downarrow \\ \mathrm{Exal}_X(V, N) & \longrightarrow & \mathrm{O}^1(V, N) \end{array} \qquad \begin{array}{ccc} \ker \mathrm{o}^{l-1}(T, g_*N) & \longrightarrow & \mathrm{O}^l(T, g_*N) \\ \downarrow & & \downarrow \\ \ker \mathrm{o}^{l-1}(V, N) & \longrightarrow & \mathrm{O}^l(V, N) \end{array}$$

Here the leftmost map is the map ψ of Lemma 1.8 (5). We also require for every open immersion (resp. étale morphism) of affine schemes $h\colon W \to V$, an isomorphism of functors

$$\alpha^l_{g,h}\colon C^l_h \circ C^l_g \Rightarrow C^l_{gh}.$$

Remark 7.1 (comparison with Artin's obstruction theories). An obstruction theory in the sense of [Artin 1974, 2.6] is a 1-step bounded obstruction theory "that is functorial in the obvious sense". We take this to mean étale-functorial in the above sense. Obstruction theories are usually half-exact and functorial for every morphism, but Exal is only contravariantly functorial for étale morphisms so the condition above does not make sense for arbitrary morphisms. On the other hand, for **Aff**-homogeneous stacks, Exal is *covariantly* functorial for every affine morphism (Lemma 1.8(2)) and the minimal obstruction theory Obs_Φ is étale-functorial (Lemma 1.8(5)).

We have the following simple lemma:

Lemma 7.2. *Let S be a locally noetherian scheme and let $\Phi\colon Y \to Z$ be a 1-morphism of* **Nil***-homogeneous S-groupoids. Let $\{\mathrm{o}^l, \mathrm{O}^l\}_{l=1}^n$ be an n-step relative obstruction theory for Φ. Let $\widetilde{\mathrm{O}}^l(T, M) \subseteq \mathrm{O}^l(T, M)$ be the image of $\mathrm{o}^l(T, M)$ for $l = 1, \ldots, n$. Then $\{\mathrm{o}^l, \widetilde{\mathrm{O}}^l\}_{l=1}^n$ is an n-step relative obstruction theory for Φ. Moreover, let $\mathrm{Obs}^l(T, -) = \mathrm{Exal}_Z(T, -)/\ker \mathrm{o}^l$ and $\mathrm{Obs}^0(T, -) = 0$. Then $\mathrm{Obs}^n(T, -) = \mathrm{Obs}_\Phi(T, -)$ and we have exact sequences*

$$0 \longrightarrow \widetilde{\mathrm{O}}^l(T, -) \longrightarrow \mathrm{Obs}^l(T, -) \longrightarrow \mathrm{Obs}^{l-1}(T, -) \longrightarrow 0$$

for $l = 1, 2, \ldots, n$. In particular, if the obstruction theory is (weakly) bounded, then so is the minimal obstruction theory $\mathrm{Obs}_\Phi(T, -)$.

We now introduce variations of Conditions 6.3(iii) and 7.3(iii) (constructible and Zariski local obstructions) in terms of an n-step relative obstruction theory.

Condition 7.3 (constructible obstructions II). There exists a weakly bounded n-step relative obstruction theory for Φ, $\{\mathrm{o}^l(-,-), \mathrm{O}^l(-,-)\}_{l=1}^n$, such that for every affine irreducible Y-scheme T that is locally of finite type over S, the obstruction spaces $\mathrm{O}^l(T, -)|_{T_0}\colon \mathsf{QCoh}(T_0) \to \mathsf{Ab}$, are GI for $l = 1, \ldots, n$ where $T_0 = T_{\mathrm{red}}$.

Condition 7.4 (Zariski local obstructions II). There exists a functorial, n-step relative obstruction theory for Φ, $\{\mathrm{o}^l(-,-), \mathrm{O}^l(-,-)\}_{l=1}^n$, such that for every affine irreducible Y-scheme T that is locally of finite type over S and whose generic point $\eta \in |T|$ is of finite type, and for every open subscheme $U \subseteq T$, the canonical maps $\mathrm{O}^l(T, \kappa(\eta)) \to \mathrm{O}^l(U, \kappa(\eta))$ are injective for $l = 1, \ldots, n$.

Lemma 7.5. *Let S be a locally noetherian scheme and let $\Phi\colon Y \to Z$ be a* 1*-morphism of* **Nil***-homogeneous S-groupoids.*

(1) (*Constructibility*) *Φ has bounded and constructible obstructions* (Conditions 6.1(iii) *and* 6.3(iii)) *if and only if Φ satisfies* Condition 7.3.

(2) (*Zariski localization*) *Φ has Zariski local obstructions* (Condition 6.4(iii)) *if and only if Φ satisfies* Condition 7.4.

Proof. If Φ has bounded deformations and obstructions (Conditions 6.1(iii) and 6.3(iii)), then the minimal obstruction theory satisfies Condition 7.3. Conversely, assume that we are given an obstruction theory $\mathrm{O}^l(-,-)$ as in Condition 7.3. Let T be an affine irreducible Y-scheme that is locally of finite type over S. Then the subfunctors $\widetilde{\mathrm{O}}^l(T,-)|_{T_0} \subseteq \mathrm{O}^l(T,-)|_{T_0}$ of Lemma 7.2 are also GI and weakly bounded by Lemma 3.10(4). Since $\mathrm{Obs}_\Phi(T,-)$ is an iterated extension of the $\widetilde{\mathrm{O}}^l(T,-)$'s, it follows that $\mathrm{Obs}_\Phi(T,-)|_{T_0}$ is GI and weakly bounded by Lemma 3.10(5b) — thus Φ has bounded and constructible obstructions (Conditions 6.1(iii) and 6.3(iii)).

If Condition 6.4(iii) holds, then the minimal obstruction theory satisfies 7.4. That Condition 7.4 implies Condition 6.4(iii) follows from Lemma 7.2. □

8. Conditions on obstructions without an obstruction theory

In this section we give conditions without reference to linear obstruction theories, just as in [Artin 1969b, Theorem 5.3 [5′c]; Starr 2006]. In the comparison we provide between our conditions on obstructions we use **Aff**-homogeneity, while Artin uses **DVR**-homogeneity and Starr uses homogeneity along localization morphisms (not just Zariski localizations). Starr's localization-homogeneity is stronger than **DVR**-homogeneity, but weaker than **Aff**-homogeneity. In Lemma 10.4, however, we establish that **DVR**-homogeneity implies **Aff**-homogeneity in all cases relevant to the proof of the Main Theorem.

Definition 8.1 [Artin 1969b, 5.1; Starr 2006, Definition 2.1]. By a *deformation situation* for $\Phi\colon Y \to Z$, we will mean data $(T \hookrightarrow T', M)$, where T is an irreducible affine Y-scheme that is locally of finite type over S, where M is a quasicoherent $\mathcal{O}_{T_{\mathrm{red}}}$-module, and where $T \hookrightarrow T'$ is an Z-extension of T by M. We say that the deformation situation is *obstructed* if the Z-extension $T \hookrightarrow T'$ cannot be lifted to a Y-extension $T \hookrightarrow T'$.

Notation 8.2. For a deformation situation $(T \hookrightarrow T', M)$, let $T_0 = T_{\mathrm{red}}$, let $\eta_0 = \operatorname{Spec} K_0$ denote the generic point of T_0, let $\eta = \operatorname{Spec}(\mathcal{O}_{T,\eta_0})$, and let $\eta' = \operatorname{Spec}(\mathcal{O}_{T',\eta_0})$. Thus $\eta \hookrightarrow \eta'$ is a Z-extension of η by $M_\eta = M \otimes_{\mathcal{O}_{T_0}} K_0$.

Condition 8.3 (constructible obstructions III). Given a deformation situation such that M is a free $\mathcal{O}_{T_0}$-module of finite rank and such that for every nonzero $\mathcal{O}_{T_0}$-module map $\epsilon\colon M_\eta \to K_0$, the resulting Z-extension $\eta \hookrightarrow \eta'_\epsilon$ of η by K_0 is obstructed, then there exists a dense open subset $U_0 \subseteq |T_0|$ such that for all points $u \in U_0$ of finite type, and all nonzero $\mathcal{O}_{T_0}$-module maps $\gamma\colon M \to \kappa(u)$, the resulting Z-extension $T \hookrightarrow T'_\gamma$ of T by $\kappa(u)$ is obstructed.

Lemma 8.4. *Let S be a locally noetherian scheme and let $\Phi\colon Y \to Z$ be a 1-morphism of limit preserving,* **Aff***-homogeneous S-groupoids. If Φ has bounded obstructions* (Condition 6.1(iii)), *then Φ has constructible obstructions* (Condition 6.3(iii)) *if and only if Φ satisfies* Condition 8.3.

Proof. Fix an irreducible affine Y-scheme T and let T_0 be its reduction. To see that Conditions 6.3(iii) and 8.3 are equivalent we will use condition (†) of Proposition 3.11 for $F(-) = \mathrm{Obs}_\Phi(T, -)|_{T_0}$. Some care is needed, though, as these two conditions are not quite equivalent for a fixed T.

Consider a deformation situation $(T \hookrightarrow T', M)$ as in Condition 8.3 and let $\omega \in F(M) = \mathrm{Obs}_\Phi(T, M)$ be the obstruction of the deformation situation. Then for every nonzero $\epsilon\colon M \to K_0$, the element $\epsilon_*\omega \in F(K_0)$ is nonzero since its image under $F(K_0) = \mathrm{Obs}_\Phi(T, K_0) \to \mathrm{Obs}_\Phi(T_\eta, K_0)$ is nonzero. If F is GI, then condition (†) is satisfied for F, M and ω. Thus, there is an open dense subset $U_0 \subseteq |T_0|$ such that $\gamma_*\omega \in F(\kappa(u))$ is nonzero for all $u \in U_0$ of finite type and nonzero maps $\gamma\colon M \to \kappa(u)$, that is, Condition 8.3 holds.

Conversely, let f, M and ω be as in condition (†) for $F(-)$. Let $V_0 = \mathrm{Spec}(A_f) \subseteq T_0 = \mathrm{Spec}\, A$ and let $V \subseteq T$ denote the corresponding open subscheme. Since Y and Z are **Aff**-homogeneous, the natural morphism $F(-)|_{A_f} = \mathrm{Obs}_\Phi(T, -)|_{V_0} \to \mathrm{Obs}_\Phi(V, -)|_{V_0}$ is an isomorphism (Lemma 1.8(5)). Since M is an A_f-module, we may thus consider $\omega \in F(M)$ as an obstruction class in $\mathrm{Obs}_\Phi(V, M)$. This class can be realized by a deformation situation $(V \hookrightarrow V', M)$. We assume that Condition 8.3 holds for this deformation situation.

Since Y and Z are **Aff**-homogeneous, we also have an isomorphism $\mathrm{Obs}_\Phi(T, -)|_{\eta_0} \to \mathrm{Obs}_\Phi(\eta, -)|_{\eta_0}$. In particular, for all $\epsilon\colon M_\eta \to K_0$, the resulting Z-extension $\eta \hookrightarrow \eta'_\epsilon$ of η by K_0 is obstructed. Thus, there exists a dense open subset $U_0 \subseteq |V_0|$ such that for all points $u \in U_0$ of finite type and maps $\gamma\colon M \to \kappa(u)$, the induced Z-extension $(V \hookrightarrow V'_\gamma, \kappa(u))$ is obstructed. In particular, $\gamma_*\omega \in F(\kappa(u)) = \mathrm{Obs}_\Phi(T, \kappa(u)) = \mathrm{Obs}_\Phi(V, \kappa(u))$ is nonzero. Thus, condition (†) holds for the given f, M and ω with $V_\omega = U_0$.

Thus, if for a given T, Condition 8.3 holds for all deformation situations $(V_0 \hookrightarrow V, M)$ where $V \subseteq T$ is an open subscheme, then F is GI. □

Remark 8.5. If S is of finite type over a Dedekind domain as in [Artin 1969b] (or Jacobson), then in Condition 8.3 it is enough to consider closed points $u \in U$. Indeed, in the proof of the lemma above, we are free to pass to open dense subsets and every S-scheme of finite type has a dense open subscheme which is Jacobson.

9. Effectivity

We begin with the following definition:

Definition 9.1. Let X be a category fibered in groupoids over the category of S-schemes. We say that X is *weakly effective* (resp. *effective*) if for every local noetherian ring $(B, \mathfrak{m})$, such that B is $\mathfrak{m}$-adically complete, with an S-scheme structure $\mathrm{Spec}\, B \to S$ such that the induced morphism $\mathrm{Spec}(B/\mathfrak{m}) \to S$ is locally of finite type, the natural functor:

$$X(\mathrm{Spec}\, B) \to \varprojlim_n X(\mathrm{Spec}(B/\mathfrak{m}^{n+1}))$$

is dense and fully faithful (resp. an equivalence). Here dense means that for every object $(\xi_n)_{n\geq 0}$ in the limit and for every $k \geq 0$, there exists an object $\xi \in X(\operatorname{Spec} B)$ such that its image in $X(\operatorname{Spec}(B/\mathfrak{m}^{k+1}))$ is isomorphic to ξ_k.

If X is an algebraic stack, then the functor $X(\operatorname{Spec} B) \to \varprojlim_n X(\operatorname{Spec}(B/\mathfrak{m}^{n+1}))$ is an equivalence of categories — thus every algebraic stack is effective. Also, it is clear that effectivity implies weak effectivity. We will see in Proposition 9.3 that the converse holds under mild hypotheses.

The following lemma is well-known, with the difficult parts attributed to Schlessinger [1968] and Rim [SGA 7_{I} 1972, Exposé VI].

Lemma 9.2. *Let S be a noetherian scheme and let X be an S-groupoid. Let $\operatorname{Spec}\Bbbk$ be an X-scheme, locally of finite type over S, such that $\Bbbk$ is a field. If X is*

(1) $\mathbf{Art}^{\mathrm{triv}}$*-homogeneous*,

(2) *weakly effective, and*

(3) *has bounded deformations at* $\operatorname{Spec}\Bbbk$ (Condition 6.1(ii)),

then there exists a pointed and affine X-scheme (T, t) such that

(a) *the point $t \in |T|$ is closed and the X-schemes* $\operatorname{Spec}\Bbbk$ *and* $\operatorname{Spec}\kappa(t)$ *are isomorphic*;

(b) *the X-scheme T is formally versal at $t \in |T|$; and*

(c) *T is affine, local, noetherian, and complete.*

Proof. By Schlessinger–Rim (e.g., [Stacks Project, Tag 06IW]), there exists an affine, local, noetherian, and complete scheme $(T = \operatorname{Spec} R, \mathfrak{m})$ and an object $(\eta_n)_{n\geq 0} \in \varprojlim_n X(T_n)$, where $T_n = \operatorname{Spec}(R/\mathfrak{m}^{n+1})$, which is a formally versal deformation (in the sense of Schlessinger–Rim) of the X-scheme structure on $\operatorname{Spec}\Bbbk$. Since X is weakly effective, there exists $\xi \in X(T)$ such that $\xi|_{T_1} \simeq \eta_1$ in $X(T_1)$. By formal versality, there exists a map of S-schemes $\phi\colon T \to T$ which restricts to the identity map on T_1 and such that $\xi|_{T_n} \simeq \phi^*\eta_n$ for every n. It is well-known that the first condition implies that ϕ is an isomorphism, hence ξ is formally versal. □

We now have the main result of this section.

Proposition 9.3. *Let S be a noetherian scheme. Let X be an S-groupoid that is*

(1) $\mathbf{Art}^{\mathrm{triv}}$*-homogeneous*,

(2) *weakly effective, and*

(3) *has bounded deformations at every X-scheme* $\operatorname{Spec}\Bbbk$, *locally of finite type over S, such that $\Bbbk$ is a field* (Condition 6.1(ii)).

Let $(B, \mathfrak{m})$ be a local noetherian ring, complete with respect to its $\mathfrak{m}$-adic topology, such that $\operatorname{Spec}(B/\mathfrak{m}) \to S$ *is locally of finite type. If $\{J_n\}_{n\geq 0}$ is an $\mathfrak{m}$-stable filtration of B (e.g., $J_n = \mathfrak{m}^{n+1}$), then the natural functor*

$$X(\operatorname{Spec} B) \to \varprojlim_n X(\operatorname{Spec}(B/J_n))$$

is an equivalence. In particular, X is effective.

Proof. Since $\mathfrak{m}$-stable filtrations of B have bounded difference [Atiyah and Macdonald 1969, Lemma 10.6] (in particular, there exists an n_0 such that $J_{n+n_0} \subseteq \mathfrak{m}^{n+1}$ for all $n \geq 0$), it is sufficient to prove the result when $J_n = \mathfrak{m}^{n+1}$. In this case, the functor above is already assumed to be fully faithful; thus, it remains to establish that it is essentially surjective. To see this, let $(\xi_n)_{n\geq 0} \in \varprojlim_n X(\operatorname{Spec}(B/\mathfrak{m}^{n+1}))$. Now apply Lemma 9.2 to the X-scheme structure on $\operatorname{Spec}(B/\mathfrak{m})$ determined by ξ_0. This produces an affine, local, noetherian, and complete X-scheme T, formally versal at its closed point t, such that the X-schemes $\operatorname{Spec}\kappa(t)$ and ξ_0 are isomorphic. By formal versality, there exists a compatible system of maps $b_n\colon \operatorname{Spec}(B/\mathfrak{m}^{n+1}) \to T$ lifting the X-scheme structures ξ_n. It follows that there is an induced map of schemes $\operatorname{Spec} B \to T$ which, by construction, defines an object $\xi \in X(\operatorname{Spec} B)$ with image $(\xi_n)_{n\geq 0} \in \varprojlim_n X(\operatorname{Spec} B/\mathfrak{m}^{n+1})$. The result follows. □

10. Proof of Main Theorem

In this section, we prove the Main Theorem. Before we do this, however, there are several preliminary results that we must prove. Conrad and de Jong [2002, Theorem 1.5] extended Artin's algebraization theorem [1969b, Theorem 1.6] to excellent rings. The following lemma summarizes their result in the language of this paper.

Theorem 10.1. *Let S be an excellent scheme and let X be an S-groupoid. Let* $\operatorname{Spec}\Bbbk$ *be an X-scheme, locally of finite type over S, such that $\Bbbk$ is a field. If X is*

(1) *limit preserving*,

(2) *weakly effective*,

(3) $\mathbf{Art}^{\mathrm{triv}}$*-homogeneous, and*

(4) *has bounded deformations at* $\operatorname{Spec}\Bbbk$ (Condition 6.1(ii)),

then there exists a pointed and affine X-scheme (T, t) such that

(a) *T is locally of finite type over S*;

(b) *the point $t \in |T|$ is closed and the X-schemes* $\operatorname{Spec}\Bbbk$ *and* $\operatorname{Spec}\kappa(t)$ *are isomorphic*; *and*

(c) *the X-scheme T is formally versal at $t \in |T|$.*

We now obtain the following algebraicity criterion for groupoids.

Proposition 10.2. *Let S be an excellent scheme. An S-groupoid X is an algebraic S-stack, locally of finite presentation over S, if and only if*

(1) *X is a stack over* $(\mathsf{Sch}/S)_{\acute{E}t}$;

(2) *X is limit preserving*;

(3) *X is weakly effective*;

(4) *X is* $\mathbf{Art}^{\mathrm{insep}}$*-homogeneous*;

(5) *X is* **rCl***-homogeneous*;

(6a) *X has bounded deformations* (Condition 6.1(ii));

(6b) *X has constructible extensions* (Condition 4.2);

(6c) *X has Zariski local extensions* (Condition 2.11); *and*

(7) *the diagonal morphism $\Delta_{X/S}\colon X \to X \times_S X$ is representable by algebraic spaces.*

Proof. The hypotheses imply that for every pair $(\operatorname{Spec}\Bbbk \xrightarrow{x} S, \xi)$, where $\Bbbk$ is a field, x is a morphism locally of finite type, and $\xi \in X(x)$, there exists a pointed and affine X-scheme (T_ξ, t) as in Theorem 10.1. Condition (7) implies that $T_\xi \to X$ is representable by algebraic spaces.

As X is **rCl**-homogeneous and has bounded deformations (Condition 6.1(ii)), Lemma 6.2 implies that X has bounded extensions (Condition 4.1). Also by Lemma 1.9, X is $\mathbf{Art}^{\mathrm{fin}}$-homogeneous. Since X has Zariski local, bounded and constructible extensions (Conditions 2.11, 4.1, and 4.2), it follows from Theorem 4.4 that we are free to assume — by passing to an affine open neighborhood of t — that the X-scheme T_ξ is formally smooth.

We finish the proof in the same manner as the proof of [Hall 2017, Theorem 7.1]: define K to be the set of all morphisms $x\colon \operatorname{Spec}\Bbbk \to S$ that are locally of finite type, where $\Bbbk$ is a field. Set $T = \coprod_{x\in K, \xi\in X(x)} T_\xi$. Then the X-scheme T is representable by smooth morphisms of algebraic spaces. We will be done if we can prove that it is representable by surjective morphisms of algebraic spaces. Since X is limit preserving, this assertion may be verified on affine X-schemes V of finite type over S. By construction, the image of the morphism $T \times_X V \to V$ contains all points of finite type; since the morphism is smooth, this image is also open. The result follows. □

The following bootstrap result will be applied several times in this section.

Lemma 10.3. *Let S be a scheme and let X be an S-groupoid. Let W be an $X \times_S X$-scheme. Let $(\Delta_{X/S})_W\colon D_{X/S,W} \to W$ be the W-groupoid obtained as the pull-back of $\Delta_{X/S}\colon X \to X \times_S X$ along W. This is equivalent to a presheaf on Sch/W.*

(1) *Let $P \subseteq$* **Aff** *be a class of morphisms and let T be a $D_{X/S,W}$-scheme. If $X \to S$ is P-homogeneous at T, then $D_{X/S,W} \to W$ is P-homogeneous at T. In particular, if $X \to S$ is P-homogeneous, then $D_{X/S,W} \to W$ is P-homogeneous.*

(2) *Let T be a $D_{X/S,W}$-scheme. If $X \to S$ is* **Nil***-homogeneous at T, then $D_{X/S,W} \to W$ is* **Nil***-homogeneous at T and there are natural isomorphisms for every quasicoherent $\mathcal{O}_T$-module M:*

$$\operatorname{Aut}_{(\Delta_{X/S})_W}(T, M) \cong 0, \quad \operatorname{Def}_{(\Delta_{X/S})_W}(T, M) \cong \operatorname{Aut}_{X/S}(T, M), \quad \operatorname{Obs}_{(\Delta_{X/S})_W}(T, M) \subseteq \operatorname{Def}_{X/S}(T, M).$$

(3) *If X is a stack over $(\mathsf{Sch}/S)_{\acute{E}t}$ (resp. $(\mathsf{Sch}/S)_{\mathrm{fppf}}$), then $D_{X/S,W}$ is a sheaf over $(\mathsf{Sch}/W)_{\acute{E}t}$ (resp. $(\mathsf{Sch}/W)_{\mathrm{fppf}}$).*

(4) *If X is limit preserving over S, then $D_{X/S,W}$ is limit preserving over W.*

(5) *If S is noetherian, W is locally of finite type over S and X is effective over S, then $D_{X/S,W}$ is effective over W.*

Proof. For (1), if $X \to S$ is P-homogeneous at T, then so is $X \times_S X$ and $\Delta_{X/S}$ [Hall 2017, Lemma 1.5(5,7,8)]. Thus, $D_{X/S,W} \to W$ is P-homogeneous at T [Hall 2017, Lemma 1.5(6)]. The assertion (2) follows from (1) and [loc. cit., Corollary 6.14]. The assertions (3) and (5) are straightforward. Finally, (4) follows from [loc. cit., Lemma 3.2(5,6)]. □

In the following lemma, we establish that under very weak boundedness hypotheses, homogeneity at artinian schemes is sufficient to imply many other forms of homogeneity.

Lemma 10.4. *Let S be an excellent scheme. Let X be an S-groupoid that is*

(1) *a stack over* $(\mathsf{Sch}/S)_{\acute{E}t}$;

(2) *limit preserving*;

(3) *weakly effective*;

(4) $\mathbf{Art}^{\mathrm{triv}}$*-homogeneous*; *and*

(5) *has bounded automorphisms and deformations at every X-scheme* $\operatorname{Spec} \Bbbk$, *locally of finite type over S, such that $\Bbbk$ is a field* (Conditions 6.1(i), (ii)).

The following assertions hold:

(a) *X is effective.*

(b) *X is* **rCl***-homogeneous.*

(c) *If X is* $\mathbf{Art}^{\mathrm{fin}}$*-homogeneous, then X is* **Int***-homogeneous.*

(d) *If X is* $\mathbf{Art}^{\mathrm{fin}}$*-homogeneous and* **DVR***-homogeneous and $\Delta_{X/S}\colon X \to X \times_S X$ is representable by algebraic spaces, then X is* **Aff***-homogeneous.*

Proof. That X is effective is Proposition 9.3. We first establish that if X satisfies the conditions (1)–(5) and $(\mathrm{H}_1^{\mathbf{rCl}})$ (resp. $(\mathrm{H}_1^{\mathbf{Int}})$), then assertion (b) (resp. (c)) holds. Fix an **rCl**-nil (resp. **Fin**-nil) pair $(\operatorname{Spec} A \to \operatorname{Spec} B, \operatorname{Spec} A \to \operatorname{Spec} A')$ such that B is the completion of an $\mathcal{O}_S$-algebra B_0 of finite type at a maximal ideal $\mathfrak{m}_0$ and $A' \to A$ and $B \to A$ are of finite type. By Lemma B.3(5), it is sufficient to prove that the functor

$$X(\operatorname{Spec} B') \to X(\operatorname{Spec} B) \times_{X(\operatorname{Spec} A)} X(\operatorname{Spec} A')$$

is essentially surjective, where $B' = B \times_A A'$. Since A is complete and $B \to A$ is finite, $A = \prod_{i=1}^n A_i$ in the category of B-algebras, where each A_i is a finite and local B-algebra. Arguing as in the proof of Lemma 1.9, we may thus reduce to the situation where A and A' are local.

Since $B' \to B$ is surjective with nilpotent kernel and B is local, B' is local with maximal ideal $\mathfrak{m}'$. For each integer $n \geq 0$ let $B'_n = B'/\mathfrak{m}'^{n+1}$, $B_n = B \otimes_{B'} B'_n$, $A_n = A \otimes_{B'} B'_n$ and $A'_n = A' \otimes_{B'} B'_n$. The pair $(\operatorname{Spec} A_n \to \operatorname{Spec} B_n, \operatorname{Spec} A_n \to \operatorname{Spec} A'_n)$ is $\mathbf{Art}^{\mathrm{triv}}$-nil (resp. $\mathbf{Art}^{\mathrm{fin}}$-nil). Let $C_n = B_n \times_{A_n} A'_n$. Note

that $\varprojlim_n C_n = B \times_A A' = B'$ and that for every $n \geq \ell$, the induced map $C_n/\mathfrak{m}'^{\ell+1}C_n \to C_\ell$ is surjective but not necessarily injective. Now $\mathbf{Art}^{\mathrm{triv}}$-homogeneity (resp. $\mathbf{Art}^{\mathrm{fin}}$-homogeneity) implies that

$$X(\operatorname{Spec} C_n) \to X(\operatorname{Spec} B_n) \times_{X(\operatorname{Spec} A_n)} X(\operatorname{Spec} A'_n)$$

is an equivalence. By Proposition 9.3, it follows that there is an equivalence

$$X(\operatorname{Spec} B) \times_{X(\operatorname{Spec} A)} X(\operatorname{Spec} A') \simeq \varprojlim_n (X(\operatorname{Spec} B_n) \times_{X(\operatorname{Spec} A_n)} X(\operatorname{Spec} A'_n)).$$

It remains to prove that the natural functor $X(\operatorname{Spec} B') \to \varprojlim_n X(\operatorname{Spec} C_n)$ is essentially surjective. To see this, we note that the map $B' \to C_n$ is surjective with kernel $K_n = B' \cap \mathfrak{m}^n(B \oplus A')$. By the Artin–Rees Lemma [Atiyah and Macdonald 1969, Propositon 10.9], the filtration $\{K_n\}_{n\geq 0}$ on B' is $\mathfrak{m}$-stable. By Proposition 9.3, the claim follows.

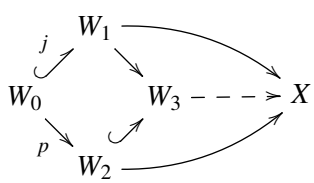
To deduce (b) (resp. (c)) in general, we apply a bootstrapping procedure. By Lemma B.2(4), to prove that X satisfies $(\mathrm{H}_1^{\mathbf{rCl}})$ (resp. $(\mathrm{H}_1^{\mathbf{Int}})$), it is sufficient to prove that $D_{X/S,W}$ is **rCl**-homogeneous (resp. **Int**-homogeneous) for every affine scheme W of finite type over S. Fix an affine scheme W of finite type over S. First observe that W is excellent. By Lemma 10.3, $D_{X/S,W}$ satisfies the hypotheses (1)–(5) and the hypothesis in (b) (resp. (c)). Indeed, **Nil**-homogeneity at $\operatorname{Spec}\Bbbk$ is equivalent to $\mathbf{Art}^{\mathrm{triv}}$-homogeneity at $\operatorname{Spec}\Bbbk$. Thus it is sufficient to prove the Lemma under the additional assumption that the diagonal of $X \to S$ is a monomorphism. Repeating this process, we see that it is sufficient to prove the Lemma when $X \to S$ is a monomorphism. In this case, however, the diagonal of $X \to S$ is an isomorphism, thus is representable and consequently satisfies $(\mathrm{H}_1^{\mathbf{Aff}})$. The claim follows.

To establish (d), we note that since X has diagonal representable by algebraic spaces, X satisfies $(\mathrm{H}_1^{\mathbf{Aff}})$. By Lemma B.3(5), it is thus sufficient to prove that

$$X(\operatorname{Spec} A_3) \to X(\operatorname{Spec} A_2) \times_{X(\operatorname{Spec} A_0)} X(\operatorname{Spec} A_1)$$

is essentially surjective for every **Aff**-nil pair $(\operatorname{Spec} A_0 \to \operatorname{Spec} A_2, \operatorname{Spec} A_0 \to \operatorname{Spec} A_1)$, where A_2 is the henselization of a finite type $\mathcal{O}_S$-algebra B at a maximal ideal $\mathfrak{m}$ and $A_2 \to A_0$ and $A_1 \to A_0$ are of finite type and $A_3 = A_2 \times_{A_0} A_1$.

Fix $(a_2, a_1, \alpha) \in X(\operatorname{Spec} A_2) \times_{X(\operatorname{Spec} A_0)} X(\operatorname{Spec} A_1)$, which we may regard as a diagram of X-schemes

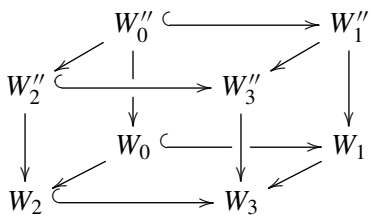


where $W_i = \operatorname{Spec} A_i$, that we must complete. Let $\Bbbk = A_2/\mathfrak{m}$; then $\operatorname{Spec}\Bbbk$ inherits an X-scheme structure from $\operatorname{Spec} A_2$. Now apply Theorem 10.1 to the X-scheme $\operatorname{Spec}\Bbbk$, which produces a pointed affine X-scheme (T, t), locally of finite type over S, which is formally versal at the closed point t. Let $W'_i = W_i \times_X T$ for $i = 0, 1, 2$ and let $p' \colon W'_0 \to W'_2$ be the pullback of $p \colon W_0 \to W_2$. Since X has

diagonal representable by algebraic spaces, W_i' is an algebraic space, locally of finite type over W_i, for each i. By construction, the morphism $W_2' \to W_2$ even admits a section $s_2 \colon W_2 \to W_2'$.

For $i = 0, 1, 2$ let $W_i'^{\mathrm{sm}} \subseteq W_i'$ denote the smooth locus of $W_i' \to W_i$, which is an open subset. By Lemma 2.2, T is formally smooth at t. Since X is **DVR**-homogeneous, T is formally smooth at every generization $t' \in |T|$ of t (Lemma 2.15). Thus $W_i'^{\mathrm{sm}}$ contains the preimage of $\mathrm{Spec}(\mathcal{O}_{T,t})$ under $W_i' \to T$. Let $Z_2 = p'(W_0' \setminus j'^{-1}(W_1'^{\mathrm{sm}}))$, $W_2'' = W_2'^{\mathrm{sm}} \setminus \overline{Z_2}$, $W_0'' = p'^{-1}(W_2'')$ and $W_1'' = j'(W_0'')$, which we regard as open subsets of $W_i'^{\mathrm{sm}}$. We claim that the section $s_2 \colon W_2 \to W_2'$ factors through W_2''. To see this, it is sufficient to check that $\overline{Z_2}$ does not contain any points above t. But Z_2 does not contain any points above $\mathrm{Spec}(\mathcal{O}_{T,t})$ and since every point of $\overline{Z_2}$ is a specialization of a point in Z_2, the claim follows.

By restriction, there is an induced section $s_0 \colon W_0 \to W_0''$. Since $W_1'' \to W_1$ is smooth and W_0 is affine, the section s_0 lifts to a section $s_1 \colon W_1 \to W_1''$ of $W_1'' \to W_1$. By [Hall 2017, Lemma A.4], there is a commutative diagram of S-schemes



where all faces of the cube are cartesian, the top and bottom faces are cocartesian, and the map $W_3'' \to W_3$ is flat. Since the top square is cocartesian, and there are compatible maps $W_i'' \to T$ for $i \neq 3$, there is a uniquely induced map $W_3'' \to T$. The sections s_i for $i = 0, 1, 2$ glue to a section $s_3 \colon W_3 \to W_3''$ of $W_3'' \to W_3$. Taking the composition $W_3 \to W_3'' \to T \to X$ proves the result. □

We now prove a version of the Main Theorem where we assume that the diagonal is representable.

Theorem 10.5. *Let S be an excellent scheme. Then a category X, fibered in groupoids over the category of S-schemes, Sch/S, is an algebraic stack, locally of finite presentation over S, if and only if it satisfies the conditions of the Main Theorem and*

(7) *the diagonal $\Delta_{X/S} \colon X \to X \times_S X$ is representable by algebraic spaces.*

Proof. We will use the criteria of Proposition 10.2. Clearly the conditions of limit preservation (2), weak effectivity (3), bounded deformations (6a) and diagonal representable by algebraic spaces (7) of Proposition 10.2 are satisfied. Either of the stack hypotheses — (1) or (1′) — imply the étale stack condition (1) of Proposition 10.2.

Either the $\mathbf{Art}^{\mathrm{insep}}$-homogeneity hypothesis (4′), or (1) and $\mathbf{Art}^{\mathrm{triv}}$-homogeneity (4) and Lemma 1.9, imply that X is $\mathbf{Art}^{\mathrm{fin}}$-homogeneous. By Lemma 10.4, (1) or (1′), combined with (2) and (3) and bounded automorphisms and deformations (5a), implies that X is **Int**-homogeneous. In particular, (4)–(5) of Proposition 10.2 are satisfied.

Now X has constructible obstructions, by (6b) and Lemma 7.5(1). Since X also has constructible deformations (5b), it has constructible extensions (Proposition 6.5(1)). Thus, X satisfies (6b) of Proposition 10.2.

Similarly by Lemma 7.5(2) and (6c), X has Zariski local obstructions. Since X also has Zariski local deformations (5c), Proposition 6.5(2) implies that X satisfies (6c) of Proposition 10.2.

If S is Jacobson (α), then X satisfies (6c) of Proposition 10.2 (Lemma 2.12), without assuming (5c) and (6c).

If X is **DVR**-homogeneous (β), then Lemma 10.4(d) implies that X is **Aff**-homogeneous; thus, X satisfies (6c) of Proposition 10.2 (Lemma 1.8(5)), without assuming (5c) and (6c). Moreover, Lemma 8.4 implies that (6b) may be substituted for Condition 8.3. The result follows. □

We are now ready to prove the Main Theorem.

Proof of Main Theorem. We will do a bootstrapping process, similar to the proof of [Hall 2017, Theorem A]. In this instance, however, we must be more careful because we are working with a weaker homogeneity assumption.

The hypotheses (1) and (4), or (γ), imply that X is $\mathbf{Art}^{\mathrm{fin}}$-homogeneous (Lemma 1.9). By Lemma 10.4, X is effective and **Int**-homogeneous.

Let W be an $X \times_S X$-scheme, affine and locally of finite type over S. By Lemma 10.3, the W-groupoid $(\Delta_{X/S})_W \colon D_{X/S,W} \to W$ satisfies the conditions of the Main Theorem. Let V be a $D_{X/S,W} \times_W D_{X/S,W}$-scheme, affine and locally of finite type over W. By Lemma 10.3, the V-groupoid $(\Delta_{D_{X/S,W}/W})_V \colon D_{D_{X/S,W},V} \to V$ satisfies the conditions of the Main Theorem. Note, however, that $(\Delta_{D_{X/S,W}/W})_V$ is a monomorphism, so has representable diagonal. By Theorem 10.5, $(\Delta_{D_{X/S,W}/W})_V$ is algebraic and locally of finite presentation over V, so $(\Delta_{X/S})_W$ has diagonal representable by algebraic spaces. By Theorem 10.5 again, $(\Delta_{X/S})_W$ is algebraic and locally of finite presentation over W; so X has diagonal representable by algebraic spaces. A final application of Theorem 10.5 informs us that X is algebraic and locally of finite presentation over S. □

11. Comparison with other criteria

In this section we compare our algebraicity criterion with Artin's criteria [1969b; 1974], Starr's criterion [2006], the criterion of the first author [Hall 2017], the criterion in the stacks project [Stacks Project], and Flenner's criterion for openness of versality [1981].

11.1. ***Artin's algebraicity criterion for functors.*** Artin [1969b, Theorem 5.3] assumes $[0'] =$ (1) (fppf stack), $[1'] =$ (2), (limit preserving) and $[2'] =$ (3) (effectivity). Further $[4']$(b)+$[5']$(a) is **Nil**-homogeneity for irreducible schemes, which implies (4). His $[4']$(a) + (c) is boundedness, Zariski-localization and constructibility of deformations (Conditions 6.1(ii), 6.4(ii), and 6.3(ii)). His $[5']$(c) is Condition 8.3 (constructibility of obstructions). Finally, $[5']$(b) together with $[4']$(a) and $[4']$(b) implies **DVR**-homogeneity so we are in the setting of (β). Conditions on automorphisms are of course redundant for functors. Condition $[3']$(a) is only used to assure that the resulting algebraic space is locally separated (resp. separated) and condition $[3']$(b) guarantees that it is quasiseparated. If one is willing to accept nonquasiseparated algebraic spaces, no separation assumptions are necessary.

11.2. ***Artin's algebraicity criterion for stacks.*** Let us begin with correcting two typos in the statement of [Artin 1974, Theorem 5.3]. In (1) the condition should be that (S1′,2) holds for F, not merely (S1,2), and in (2) the canonical map should be fully faithful with dense image, not merely faithful with dense image. Otherwise it is not possible to bootstrap and deduce algebraicity of the diagonal.

Artin assumes that X is a stack for the étale topology, and that X is limit preserving. He assumes (1) that the Schlessinger conditions (S1′,2) hold and boundedness of automorphisms. In our terminology, (S1′) is **rCl**-homogeneity, which implies $\mathbf{Art}^{\mathrm{triv}}$-homogeneity, our (4). The other two conditions are exactly boundedness of automorphisms and deformations (5a). Artin's condition (2) is our (3) (effectivity). Artin's condition (3) is étale localization and constructibility of automorphisms, deformations and obstructions, and compatibility with completions for automorphisms and deformations. The constructibility condition is slightly stronger than our (5b) + (6b) and the étale localization condition implies the much weaker (5c) + (6c). We do not use compatibility with completions. Finally, Artin's condition (4) implies that the double diagonal of the stack is quasicompact and this condition can be omitted if we work with stacks without separation conditions. Thus [Artin 1974, Theorem 5.3] follows from our Main Theorem, except that Artin only assumes that the groupoid is a stack in the étale topology. This is related to the issue when comparing formal versality to formal smoothness mentioned in the introduction and discussed in Remark 2.8.

Remark 11.1. That automorphisms and deformations are sufficiently compatible with completions for Artin's proof to go through actually follows from the other conditions. In fact, let A be a noetherian local ring with maximal ideal $\mathfrak{m}$, let $T = \operatorname{Spec} A$ and let $T \to X$ be given. Then the injectivity of the comparison map

$$\varphi\colon\ \mathrm{Def}_{X/S}(T, M) \otimes_A \hat{A} \to \varprojlim_n \mathrm{Def}_{X/S}(T, M/\mathfrak{m}^n M)$$

for a finitely generated A-module M follows from the boundedness of $\mathrm{Def}_{X/S}(T, -)$, see Remark 3.8. If $T \to X$ is formally versal, then φ is also surjective. Indeed, from (S1) it follows that $\mathrm{Der}_S(T, M/\mathfrak{m}^n M) \to \mathrm{Def}_{X/S}(T, M/\mathfrak{m}^n M)$ is surjective for all n, so the composition

$$\mathrm{Der}_S(T, M) \otimes_A \hat{A} \cong \varprojlim_n \mathrm{Der}_S(T, M/\mathfrak{m}^n M) \to \varprojlim_n \mathrm{Def}_{X/S}(T, M/\mathfrak{m}^n M),$$

which factors through φ, is surjective.

The variant [Starr 2006, Proposition 1.1] has the same conditions as [Artin 1974, Theorem 5.3] except that it is phrased in a relative setting. From Section 6, it is clear that our conditions can be composed. The salient point is that with **rCl**-homogeneity (or even with just (S1), i.e., **rCl**-semihomogeneity, as in [Flenner 1981]), there is always a linear minimal obstruction theory. There is further an exact sequence relating the minimal obstruction theories for the composition of two morphisms [Hall 2017, Proposition 6.13]. Thus [Starr 2006, Proposition 1.1] also follows from our main theorem.

We wish to point out that Starr proves openness of versality [2006, Theorem 2.15] using his formalism of generic extenders [loc. cit., Definition 2.7]. This is similar to our Condition 8.3 (and Artin's analogous condition in his algebraicity criterion for functors). The main difference is that he also assumes homogeneity along localizations (not just Zariski localizations), as opposed to **DVR**-homogeneity.

11.3. ***The criterion in [Hall 2017] using coherence.*** There are two differences between [Hall 2017, Theorem A] and our main theorem. The first is that Condition (4) is strengthened to **Aff**-homogeneity. As this includes **DVR**-homogeneity, (5c) and (6c) become redundant. Zariski localization also follows immediately from **Aff**-homogeneity without involving **DVR**-homogeneity, see the discussion after Condition 2.11. We thus have the following version of our Main Theorem.

Theorem 11.2. *Let S be an excellent scheme. Then a category X that is fibered in groupoids over the category of S-schemes, Sch/S, is an algebraic stack that is locally of finite presentation over S, if and only if it satisfies the following conditions*:

(1′) *X is a stack over $(\mathsf{Sch}/S)_{\acute{E}t}$.*

(2) *X is limit preserving.*

(3) *X is effective.*

(4″) *X is* **Aff***-homogeneous.*

(5a) *Automorphisms and deformations are bounded* (Conditions 6.1(i)–(ii)).

(5b) *Automorphisms and deformations are constructible* (Conditions 6.3(i)–(ii)).

(6b) *Obstructions are constructible* (Condition 6.3(iii), *or* 7.3, *or* 8.3).

The second difference is that (5a), (5b), and (6b) are replaced with the condition that $\mathrm{Aut}_{X/S}(T, -)$, $\mathrm{Def}_{X/S}(T, -)$, $\mathrm{Obs}_{X/S}(T, -)$ are *coherent* functors. This implies that the functors are bounded and CB (Example 3.6), hence satisfy (5a), (5b), and (6b).

11.4. ***The criterion in the Stacks project.*** In the Stacks project, the basic version of Artin's axiom [Stacks Project, Tags 07XJ, 07Y5] requires that

[0] X is a stack in the étale topology;

[1] X is limit preserving;

[2] X is $\mathbf{Art}^{\mathrm{fin}}$-homogeneous (this is the Rim–Schlessinger condition RS);

[3] $\mathrm{Aut}_{X/S}(\operatorname{Spec} k, k)$ and $\mathrm{Def}_{X/S}(\operatorname{Spec} k, k)$ are finite dimensional;

[4] X is effective;

[5] X, Δ_X and Δ_{Δ_X} satisfy openness of versality.

There is also a criterion for when X satisfies openness of versality [Stacks Project, Tag 07YU] using naive obstruction theories with finitely generated cohomology groups. This uses the (RS*)-condition which is our **Aff**-homogeneity [Stacks Project, Tag 07Y8]. The existence of the naive obstruction theory implies

that $\mathrm{Aut}_{X/S}(T,-)$, $\mathrm{Def}_{X/S}(T,-)$, $\mathrm{Obs}_{X/S}(T,-)$ are bounded and CB (Example 3.4), hence satisfy (5a), (5b), and (6b) when T is an affine X-scheme that is locally of finite type over S.

In [Stacks Project], the condition that the base scheme S is excellent is replaced with the condition that its local rings are G-rings. In our treatment, excellency enters at two places: in the application of Néron–Popescu desingularization in Proposition 10.2 via [Conrad and de Jong 2002] and in the context of **DVR**-homogeneity in Lemma 2.15. In both cases, excellency can be replaced with the condition that the local rings are G-rings without modifying the proofs.

11.5. ***Flenner's criterion for openness of versality.*** Flenner [1981] does not give a precise analogue of our main theorem, but his main result (Satz 4.3) is a criterion for the openness of versality. In his criterion he has a limit preserving S-groupoid which satisfies (S1)–(S4). The first condition (S1) is identical to Artin's condition (S1), i.e., **rCl**-semihomogeneity. The second condition (S2) is boundedness and Zariski localization of deformations. The third condition (S3) is boundedness and Zariski localization of the minimal obstruction theory. Finally (S4) is constructibility of deformations and obstructions. The Zariski localization condition is incorporated in the formulation of (S3) and (S4) which deals with *sheaves* of deformation and obstructions modules. His (S2)–(S4) are marginally stronger than our conditions, for example, treating arbitrary schemes instead of irreducible schemes. Satz 4.3 [Flenner 1981] thus becomes the first part of Theorem 4.4, in view of Section 6, except that we assume **rCl**-homogeneity instead of **rCl**-semihomogeneity. This is a pragmatic choice that simplifies matters since $\mathrm{Exal}_X(T,M)$ becomes a module instead of a pointed set. Also, in any algebraicity criterion, we would need homogeneity to deduce that the diagonal is algebraic and, conversely, if the diagonal is algebraic, then semihomogeneity implies homogeneity.

11.6. ***Criterion for local constructibility.*** There is a useful criterion for when a sheaf (or a stack) is locally constructible, that is, when it corresponds to an étale algebraic space (or algebraic stack) [Artin 1973, Chapitre VII, Théorème 7.2]:

Theorem 11.3. *Let S be an excellent scheme. Then a category X that is fibered in groupoids over Sch/S, is an algebraic stack that is étale over S, if and only if it satisfies the following conditions*:

(1) *X is a stack over $(\mathsf{Sch}/S)_{Ét}$.*

(2) *X is limit preserving.*

(3) *$X(B)\to X(B/\mathfrak{m})$ is an equivalence of categories for every local noetherian ring $(B,\mathfrak{m})$, such that B is $\mathfrak{m}$-adically complete, with an S-scheme structure $\operatorname{Spec} B\to S$ such that the induced morphism $\operatorname{Spec}(B/\mathfrak{m})\to S$ is of finite type.*

The necessity of the conditions is clear. That the conditions are sufficient can be proven directly as follows. Let $j\colon (\mathsf{Sch}/S)_{\text{Ét}}\to S_{\text{ét}}$ denote the morphism of topoi corresponding to the inclusion of the small étale site into the big étale site. It is enough to prove that $j^{-1}j_*X\to X$ is an equivalence. As X is limit preserving, it is enough to verify that $f^*(X|_{S_{\text{ét}}})\to X|_{T_{\text{ét}}}$ is an equivalence for every morphism $f\colon T\to S$

locally of finite type, and this can be checked on stalks at points of finite type. Therefore, it suffices to prove that $X(B) \to X(B/\mathfrak{m})$ is an equivalence when B is the henselization of $\mathcal{O}_{T,t}$, for every $t \in |T|$ of finite type. This follows from general Néron–Popescu desingularization and the three conditions.

A proof more in the lines of this paper goes as follows: from (3) it follows that: X is $\mathbf{Art}^{\mathrm{fin}}$-homogeneous; X is effective; and $X \to S$ is formally étale at every point of finite type. In particular, $\mathrm{Aut}_{X/S}(T, N) = \mathrm{Def}_{X/S}(T, N) = \mathrm{Obs}_{X/S}(T, N) = 0$ for every X-scheme T that is of finite type over S and every quasicoherent $\mathcal{O}_T$-module N with support that is artinian (use Lemmas 5.1 and 5.4). Thus, $\mathrm{Aut}_{X/S}(T, -) = \mathrm{Def}_{X/S}(T, -) = 0$ by Theorem 3.7. Theorem 11.3 would follow from the main theorem if we also can show that $\mathrm{Obs}_{X/S}(T, -) = 0$. As we do not yet know that $\mathrm{Obs}_{X/S}(T, -)$ is half-exact, it is unclear to us how to deduce that $\mathrm{Obs}_{X/S}(T, -) = 0$ without invoking Popescu desingularization. A more elementary approach, that does not rely on the main theorem, is to note that given an X-scheme T that is locally of finite presentation over S, and a point $t \in |T|$ of finite type, then $T \to X$ is formally smooth at t if and only $T \to S$ is formally smooth at t. Thus, openness of formal smoothness for $T \to X$ follows.

Appendix A. Approximation of integral morphisms

In this appendix, we give an approximation result for integral homomorphisms of rings.

Lemma A.1. *Let A be a ring, let B be an A-algebra and let C be a B-algebra. Assume that B and C are integral A-algebras. Then there exists a filtered system $(B_\lambda \to C_\lambda)_\lambda$ of finite and finitely presented A-algebras, with direct limit $B \to C$. In addition, if $A \to B$ (resp. $B \to C$, resp. $A \to C$) has one of the properties*:

(1) *surjective*,

(2) *surjective with nilpotent kernel*,

then the system can be chosen such that the morphisms $A \to B_\lambda$ (resp. $B_\lambda \to C_\lambda$, resp. $A \to C_\lambda$) all have the corresponding property.

If we start with a system satisfying the first part of the lemma, then it is not always the case that the second part holds after increasing λ. Therefore, the approximation $B_\lambda \to C_\lambda$ has to be built with the second part in mind.

Proof of Lemma A.1. Let Λ be the set of finite subsets of $B \amalg C$, or, if $B \to C$ is surjective, only those of B. For $\lambda = \lambda_B \cup \lambda_C \in \Lambda$, let $B^\circ_\lambda \subseteq B$ be the A-subalgebra generated by λ_B and let $C^\circ_\lambda \subseteq C$ be the A-subalgebra generated by λ_C and the image of λ_B in C.

Then $B = \varinjlim_{\lambda\in\Lambda} B^\circ_\lambda$ and $C = \varinjlim_{\lambda\in\Lambda} C^\circ_\lambda$ and we have homomorphisms $B^\circ_\lambda \to C^\circ_\lambda$ for all λ. Moreover, if $A \to B$ (resp. $B \to C$, resp. $A \to C$) is surjective or surjective with nilpotent kernel then so is $A \to B^\circ_\lambda$ (resp. $B^\circ_\lambda \to C^\circ_\lambda$, resp. $A \to C^\circ_\lambda$) for every λ.

For every λ, let $P_\lambda = A[x_i : i \in \lambda_B]$ and $Q_\lambda = A[y_j : j \in \lambda]$ be polynomial rings and let $P_\lambda \to B^\circ_\lambda$ and $Q_\lambda \to C^\circ_\lambda$ be the natural surjections. We have homomorphisms $P_\lambda \to Q_\lambda$ compatible with $B^\circ_\lambda \to C^\circ_\lambda$ and

if $B \to C$ is surjective, then $P_\lambda = Q_\lambda$. For a finite subset $L \subseteq \Lambda$, let $P_L = \bigotimes_{\lambda\in L} P_\lambda$ and $Q_L = \bigotimes_{\lambda\in L} Q_\lambda$, where the tensor products are over A.

For fixed $L \subseteq \Lambda$ choose finitely generated ideals $I_L \subseteq \ker(P_L \to B)$ and $I_L Q_L \subseteq J_L \subseteq \ker(Q_L \to C)$ and let $B_L = P_L/I_L$ and $C_L = Q_L/J_L$. If $A \to B$ (resp. $A \to C$) is surjective, then for sufficiently large I_L (resp. J_L), we have that $A \to B_L$ (resp. $A \to C_L$) is surjective. If $B \to C$ is surjective, then by construction $P_L = Q_L$ so $B_L \to C_L$ is surjective. If, in addition, $B \to C$ has nilpotent kernel with nilpotency index n, then we replace I_L with $I_L + J_L^n$ so that $B_L \to C_L$ has nilpotent kernel.

Consider the set Ξ of pairs $\xi = (L, I_L, J_L)$ where $L \subseteq \Lambda$ is a finite subset, and $I_L \subseteq P_L$ and $J_L \subseteq Q_L$ are finitely generated ideals as in the previous paragraph. Then $(B_L \to C_L)_\xi$ is a filtered system of finite and finitely presented A-algebras with direct limit $(B \to C)$ which satisfies the conditions of the lemma. □

Lemma A.2. *Let $f : X \to Y$ be a morphism of affine schemes. Let P be one of the properties* **Nil**, **Cl**, **rNil**, **rCl**, **Int**, *or* **Aff** (*see* Section 1). *If f has property P, then there exists a filtered system $(f_\lambda : X_\lambda \to Y)_\lambda$ with inverse limit $f : X \to Y$ such that every f_λ is of finite presentation with property P.*

Proof. The result is standard when $P \in \{\mathbf{Cl}, \mathbf{Nil}, \mathbf{Int}, \mathbf{Aff}\}$. For $P = \mathbf{rNil}$ (resp. $P = \mathbf{rCl}$), choose a nilpotent immersion $X_0 \to X$ such that $X_0 \to X \to Y$ is **Nil** (resp. **Cl**). The lemma then follows from Lemma A.1 with $Y = \operatorname{Spec} A$, $X = \operatorname{Spec} B$ and $X_0 = \operatorname{Spec} C$. □

Fix a scheme S and consider the category of diagrams $[Y \xleftarrow{f} X \xrightarrow{i} X']$ of S-schemes. A morphism of diagrams $\Phi : [Y_1 \xleftarrow{f_1} X_1 \xrightarrow{i_1} X_1'] \to [Y_2 \xleftarrow{f_2} X_2 \xrightarrow{i_2} X_2']$ consists of morphisms $\Phi_Y : Y_1 \to Y_2$, $\Phi_X : X_1 \to X_2$ and $\Phi_{X'} : X_1' \to X_2'$ such that the natural diagram is commutative but not necessarily cartesian. We say that Φ is affine if Φ_Y, Φ_X and $\Phi_{X'}$ are affine. Given an inverse system of diagrams with affine bonding maps, the inverse limit exists and is calculated component by component.

Proposition A.3. *Let S be an affine scheme and let P be one of the properties* **Nil**, **Cl**, **rNil**, **rCl**, **Int**, *or* **Aff**. *Let $\mathbf{W} = [Y \xleftarrow{f} X \xrightarrow{i} X']$ be a diagram of affine S-schemes where i is* **Nil**, *and f is P. Then $\mathbf{W}$ is an inverse limit of diagrams $\mathbf{W}_\lambda = [Y_\lambda \xleftarrow{f_\lambda} X_\lambda \xrightarrow{i_\lambda} X_\lambda']$ of affine finitely presented S-schemes where i_λ is* **Nil**, *and f_λ is P. Moreover, if we let $Y' = Y \sqcup_X X'$ and $Y_\lambda' = Y_\lambda \sqcup_{X_\lambda} X_\lambda'$ denote the push-outs, then $Y' = \varprojlim_{\lambda\in\Lambda} Y_\lambda'$.*

Proof. We begin by looking at the induced diagram $[Y \xrightarrow{j} Y' \xleftarrow{g} X']$. As j is a nilpotent closed immersion it follows that g has property P. We will write this diagram as an inverse limit of diagrams $[Y_\lambda \xrightarrow{j_\lambda} \overline{Y}_\lambda' \xleftarrow{g_\lambda} X_\lambda']$ of finite presentation over S where j_λ is **Nil** and g_λ has property P. To this end, we begin by writing (using Lemma A.2)

(1) $Y' = \varprojlim \overline{Y}_\alpha'$ where $\overline{Y}_\alpha' \to S$ are affine and of finite presentation;

(2) $X' = \varprojlim X_\beta'$ where $X_\beta' \to Y'$ are P and of finite presentation; and

(3) $Y = \varprojlim Y_\gamma$ where $Y_\gamma \to Y'$ are **Nil** and of finite presentation.

For every pair (β,γ) there is [EGA IV$_3$ 1966, Théorème 8.10.5] an index $\alpha_0(\beta,\gamma)$, and a cartesian diagram

$$\begin{array}{ccccc} Y_\gamma & \hookrightarrow & Y' & \hookleftarrow & X'_\beta \\ \downarrow & & \downarrow & & \downarrow \\ Y_{\alpha_0(\beta,\gamma)\beta\gamma} & \hookrightarrow & \overline{Y}'_{\alpha_0(\beta,\gamma)} & \hookleftarrow & X'_{\alpha_0(\beta,\gamma)\beta\gamma} \end{array}$$

where $X'_{\alpha_0(\beta,\gamma)\beta\gamma} \to \overline{Y}'_{\alpha_0(\beta,\gamma)}$ and $Y_{\alpha_0(\beta,\gamma)\beta\gamma} \to \overline{Y}'_{\alpha_0(\beta,\gamma)}$ are morphisms of finite presentation that are P and **Nil** respectively.

For every $\alpha \geq \alpha_0(\beta,\gamma)$ we also let $[Y_{\alpha\beta\gamma} \to \overline{Y}'_\alpha \leftarrow X'_{\alpha\beta\gamma}]$ denote the pull-back along $\overline{Y}'_\alpha \to \overline{Y}'_{\alpha_0(\beta,\gamma)}$. Let $I = \{(\beta,\gamma,\alpha)\}$ be the set of indices such that $\alpha \geq \alpha_0(\beta,\gamma)$. For every finite subset $J \subseteq I$, we let

$$\overline{Y}'_J = \prod_{(\beta,\gamma,\alpha)\in J} \overline{Y}'_\alpha, \quad Y_J = \prod_{(\beta,\gamma,\alpha)\in J} Y_{\alpha\beta\gamma}, \quad \text{and} \quad X'_J = \prod_{(\beta,\gamma,\alpha)\in J} X'_{\alpha\beta\gamma},$$

where the products are taken over S. The finite subsets $J \subseteq I$ form a partially ordered set under inclusion and the induced morphisms:

$$Y' \to \varprojlim_J \overline{Y}'_J, \quad Y \to \varprojlim_J Y_J, \quad \text{and} \quad X' \to \varprojlim_J X'_J$$

are closed immersions. Now, let $K_{Y_J} = \ker(\mathcal{O}_{Y_J} \to (g_J)_*\mathcal{O}_Y)$ and similarly for $K_{\overline{Y}'_J}$ and $K_{X'_J}$. Note that $K_{\overline{Y}'_J}\mathcal{O}_{Y_J} \subseteq K_{Y_J}$ and $K_{\overline{Y}'_J}\mathcal{O}_{X'_J} \subseteq K_{X'_J}$. We then let $\Lambda = \{(J, R_{Y_J}, R_{\overline{Y}'_J}, R_{X'_J})\}$ where $J \subseteq I$ is a finite subset and $R_{Y_J} \subseteq K_{Y_J}$, $R_{\overline{Y}'_J} \subseteq K_{\overline{Y}'_J}$ and $R_{X'_J} \subseteq K_{X'_J}$ are finitely generated ideals such that $R_{\overline{Y}'_J}\mathcal{O}_{Y_J} \subseteq R_{Y_J}$ and $R_{\overline{Y}'_J}\mathcal{O}_{X'_J} \subseteq R_{X'_J}$. For every $\lambda \in \Lambda$ we put

$$\overline{Y}'_\lambda = \operatorname{Spec}(\mathcal{O}_{\overline{Y}'_J}/R_{\overline{Y}'_J}), \quad Y_\lambda = \operatorname{Spec}(\mathcal{O}_{Y_J}/R_{Y_J}), \quad \text{and} \quad X'_\lambda = \operatorname{Spec}(\mathcal{O}_{X'_J}/R_{X'_J}).$$

Then $[Y \to Y' \leftarrow X'] = \varprojlim_\lambda [Y_\lambda \to \overline{Y}'_\lambda \leftarrow X'_\lambda]$. Finally, we take $X_\lambda = X'_\lambda \times_{\overline{Y}'_\lambda} Y_\lambda$ so that

$$[Y \xleftarrow{f} X \xrightarrow{i} X'] = \varprojlim_\lambda [Y_\lambda \xleftarrow{f_\lambda} X_\lambda \xrightarrow{i_\lambda} X'].$$

Indeed, $X = X' \times_{Y'} Y$ and inverse limits commute with fiber products.

For the last assertion, we note that all schemes are affine and that there are exact sequences

$$0 \to \Gamma(\mathcal{O}_{Y'}) \to \Gamma(\mathcal{O}_Y) \times \Gamma(\mathcal{O}_{X'}) \to \Gamma(\mathcal{O}_X) \to 0,$$
$$0 \to \Gamma(\mathcal{O}_{Y'_\lambda}) \to \Gamma(\mathcal{O}_{Y_\lambda}) \times \Gamma(\mathcal{O}_{X'_\lambda}) \to \Gamma(\mathcal{O}_{X_\lambda}) \to 0, \quad \forall \lambda \in \Lambda.$$

Note that Y'_λ can be different from $\overline{Y}'_\lambda$. As direct limits of rings are exact it follows that $Y' = \varprojlim Y'_\lambda$. □

Appendix B. Bootstrapping homogeneity

The following notation will be useful:

Notation B.1. Fix a scheme S and a 1-morphism of S-groupoids $\Phi\colon Y \to Z$.

- If W is a $Y \times_Z Y$-scheme, let $(\Delta_\Phi)_W \colon D_{\Phi,W} \to W$ denote the W-groupoid obtained by pulling back $\Delta_\Phi \colon Y \to Y \times_Z Y$ along $W \to Y \times_Z Y$.
- Fix a class P of morphisms of S-schemes. For a P-nil square over S as in (1-1), let

$$\Lambda_{Y,T'} \colon Y(T') \to Y(V') \times_{Y(V)} Y(T)$$

denote the natural functor.

The following bootstrapping lemma provides a powerful technique to verify condition (H_1^P) of Definition 1.3.

Lemma B.2. *Fix a scheme S, a class $P \subseteq \mathbf{Aff}$ of morphisms of S-schemes and a 1-morphism of S-groupoids $\Phi \colon Y \to Z$. If Z satisfies (H_1^P), then the following conditions are equivalent*:

(1) *Y satisfies (H_1^P)*;

(2) *for every geometric P-nil square over S as in (1-1), $\Lambda_{Y,T'}$ is fully faithful*;

(3) *for every $Y \times_Z Y$-scheme W, the W-groupoid $D_{\Phi,W}$ is P-homogeneous.*

In addition, if Y and Z are limit preserving Zariski stacks and P is Zariski local, then these conditions are equivalent to the following:

(4) *Φ satisfies Condition (3) for all W affine and of finite presentation over S.*

In particular, if $\Delta_{Y/S}$ is representable by algebraic spaces, then Y satisfies $(\mathrm{H}_1^{\mathbf{Aff}})$.

Condition (3) is not equivalent to P-homogeneity of Δ_Φ unless we a priori know that $Y \times_Z Y$ is P-homogeneous — an uninteresting situation.

Proof. For (1) $\Longrightarrow$ (2), fix a geometric P-nil square over S as in (1-1). We must prove that the functor $\Lambda_{Y,T'}$ is fully faithful, that is, if y_1 and y_2 are two Y-scheme structures on T' such that $\Lambda_{Y,T'}(y_1) \cong \Lambda_{Y,T'}(y_2)$, then there is a unique isomorphism of Y-schemes $y_1 \cong y_2$. Since Y satisfies (H_1^P), any Y-scheme structure on T' makes the resulting P-nil square cocartesian (because geometric P-nil squares over S are cocartesian). The claim follows.

For (2) $\Longrightarrow$ (3), we fix a $Y \times_Z Y$-scheme W. To establish (H_1^P) for $D_{\Phi,W}$, it is sufficient to prove that a geometric P-nil square over $D_{\Phi,W}$ as in (1-1) is cocartesian. There is a canonical map $T' \to W$ and this corresponds to two maps $y_1, y_2 \colon T' \to Y$ and a 2-isomorphism τ between $\Phi \circ y_1$ and $\Phi \circ y_2$. If Q is a $D_{\Phi,W}$-scheme with compatible maps from T and V', we obtain a map $T' \to Q$ over W and hence two maps $T' \to D_{\Phi,W}$. These two maps correspond to 2-isomorphisms α, β between y_1 and y_2 compatible with τ and such that $\Lambda_{Y,T'}(\alpha) = \Lambda_{Y,T'}(\beta)$. Since $\Lambda_{Y,T'}$ is faithful, we conclude that $\alpha = \beta$ and hence that the square is cocartesian over $D_{\Phi,W}$.

To establish (H_2^P) for $D_{\Phi,W}$, it is sufficient to prove that every P-nil pair over $D_{\Phi,W}$ may be completed to a P-nil square. Clearly, we can complete such a P-nil pair to a geometric P-nil square over W as in (1-1). It remains to promote T' to a $D_{\Phi,W}$-scheme. However, $T' \to W \to Y \times_Z Y$ factors through

Y because $\Lambda_{Y,T'}$ is full, $\Lambda_{Z,T'}$ is faithful, and T' comes from a P-nil pair over Y. Thus, T' lifts to a $D_{\Phi,W}$-scheme and the claim follows.

For (3) $\implies$ (1), we have to prove that a geometric P-nil square over Y as in (1-1) is cocartesian. Thus, we must prove that if Q is a Y-scheme that fits into the following P-nil square:

$$\begin{array}{ccc} V & \xrightarrow{p} & T \\ {\scriptstyle j}\downarrow & & \downarrow{\scriptstyle a} \\ V' & \xrightarrow{b} & Q \end{array}$$

then there is a unique compatible map of Y-schemes $T' \to Q$. Note that since Z satisfies (H_1^P), there is a unique Z-morphism $T' \to Q$. Thus, it is sufficient to prove that the two induced Y-scheme structures on T' coincide. So we may regard T' as a $(Y \times_Z Y)$-scheme and let $D_{\Phi,T'} \to T'$ be the pullback of $\Delta_{Y/Z}$ to T'. Since $D_{\Phi,T'}$ is P-homogeneous and $(V \to T, V \to V')$ is a P-nil pair over $D_{\Phi,T'}$, it follows that the geometric P-nil square over Y is uniquely a cocartesian P-nil square over $D_{\Phi,T'}$. The claim follows.

Noting [Hall 2017, Lemma 1.5(7)], the equivalence (4) $\iff$ (3) is routine. □

The following lemma (compare [Hall 2017, Lemma 1.5(4)]) is particularly useful when combined with Lemma B.2.

Lemma B.3. *Fix a scheme S and a limit preserving étale S-stack X. Let P be one of the properties* **Nil**, **Cl**, **rNil**, **rCl**, **Int**, *or* **Aff**. *If X satisfies (H_1^P), then the following conditions are equivalent*:

(1) *X is P-homogeneous*;

(2) *$\Lambda_{X,T'}$ is essentially surjective for every geometric P-nil square over S as in (1-1) where T, V, and V' are affine*;

(3) *Condition (2) holds when T, V, and V' are of finite presentation over S*; *or*

(4) *Condition (2) holds when T is the henselization of an affine scheme of finite presentation over S at a closed point, and $V \to T$, $V \to V'$ are of finite presentation.*

If in addition $P \subseteq$ **Int** *and S is excellent, then these conditions are equivalent to the following*:

(5) *Condition (2) holds when T' is the completion of an affine scheme of finite type over S at a closed point, and $V \to T$ is finite.*

In particular, if S is locally noetherian then condition (S1′) *of* [Artin 1974, 2.3] *is equivalent to* **rCl**-*homogeneity for X.*

Proof. Note that $\Lambda_{X,T'}$ is fully faithful (Lemma B.2) so (1) $\iff$ (2) by [Hall 2017, Lemma 1.5(4)]. Obviously, (2) $\implies$ (3), (4), and (5). To see (3) $\implies$ (2), as X is a Zariski stack we may assume that $S = \operatorname{Spec} R$ is affine. By Proposition A.3, every P-nil pair $(V \xrightarrow{p} T, V \xrightarrow{j} V')$, where T is affine, may be written as an inverse limit of P-nil pairs $(V_\lambda \xrightarrow{p_\lambda} T_\lambda, V_\lambda \xrightarrow{j_\lambda} V'_\lambda)$ of finite presentation over S such that T_λ is affine. Furthermore, T' is the inverse limit of the T'_λ, where $T'_\lambda = T_\lambda \amalg_{V_\lambda} V'_\lambda$. The assertion then follows from our assumption that X is limit preserving and [Hall 2017, Lemma 1.5(4)].

To see (4) $\Longrightarrow$ (3), we fix a geometric P-nil square over S as in (1-1) with the properties prescribed by (3). On the small flat site T'_{fl}, we can consider two fibered categories that are stacks for étale covers. The first, F_1, is just the restriction of X. The second, F_2, over a flat morphism $U' \to T'$ has fiber $X(V' \times_{T'} U') \times_{X(V \times_{T'} U')} X(T \times_{T'} U')$. The functor $F_1 \to F_2$ is fully faithful (Lemma B.2); it remains to prove that it is locally surjective. Let $t \in T$ be a closed point and let T_t^h denote the henselization of T at t. This uniquely lifts to a henselization $T_t'^h$ of T'. By assumption, $F_1(T_t'^h) \simeq F_2(T_t'^h)$. Fix $\eta \in F_2(T)$ and let η_t^h denote its image in $F_2(T_t'^h)$. It follows that there exists $\tilde{\eta}_t^h \in F_1(T_t'^h)$ inducing η_t^h. Since F_1 is limit preserving, $\tilde{\eta}_t^h$ is induced by some $\tilde{\eta}_t^{U'} \in F_1(U')$, where $(U', u) \to (T, t)$ is étale. Since $F_1 \to F_2$ is fully faithful and F_2 is limit preserving, we can arrange so that $\tilde{\eta}_t^{U'}$ agrees $\eta_t|_{U'}$. The claim follows.

Finally, to see (5) $\Longrightarrow$ (4), we will argue similarly to (4) $\Longrightarrow$ (3). So we fix a geometric P-nil square over S as in (1-1) with the properties prescribed by (4). Since $P \subseteq \mathbf{Int}$, this implies that T' is also the henselization of an affine scheme of finite type over S at a closed point; in particular, T' is excellent. Defining F_1 and F_2 analogously, we obtain a fully faithful morphism of groupoids $\phi\colon F_1 \to F_2$ over T'_{fl} which are stacks for étale covers. Let $\hat{T}'$ be the completion of T' at its unique closed point, by hypothesis we have that $F_1(\hat{T}') \simeq F_2(\hat{T}')$. Since T' is excellent, Néron–Popescu desingularization [Popescu 1986] implies that $\hat{T}'$ is an inverse limit of affine and smooth T'-schemes. Now argue just as before to deduce the claim. □

Acknowledgments

We would like to thank M. Artin for encouraging comments and L. Moret–Bailly for answering a question on MathOverflow about Jacobson schemes. We would also especially like to thank the referees for their patience, support and a number of useful comments.

References

[Angéniol 1981] B. Angéniol, *Familles de cycle algébriques—schéma de Chow*, Lecture Notes in Mathematics **896**, Springer, 1981. MR Zbl

[Artin 1969a] M. Artin, "Algebraic approximation of structures over complete local rings", *Inst. Hautes Études Sci. Publ. Math.* 36 (1969), 23–58. MR Zbl

[Artin 1969b] M. Artin, "Algebraization of formal moduli, I", pp. 21–71 in *Global Analysis (Papers in Honor of K. Kodaira)*, edited by D. C. Spencer and S. Iyanaga, Univ. Tokyo Press, 1969. MR Zbl

[Artin 1973] M. Artin, *Théorèmes de représentabilité pour les espaces algébriques*, Les Presses de l'Université de Montréal, 1973. MR Zbl

[Artin 1974] M. Artin, "Versal deformations and algebraic stacks", *Invent. Math.* **27** (1974), 165–189. MR Zbl

[Atiyah and Macdonald 1969] M. F. Atiyah and I. G. Macdonald, *Introduction to commutative algebra*, Addison-Wesley Publishing Co., Reading, Massachusetts, 1969. MR Zbl

[Auslander 1966] M. Auslander, "Coherent functors", pp. 189–231 in *Proceedings of Conference of Categorical Algebra* (La Jolla, California, 1965), edited by S. Eilenberg et al., Springer, New York, 1966. MR

[Conrad and de Jong 2002] B. Conrad and A. J. de Jong, "Approximation of versal deformations", *J. Algebra* **255**:2 (2002), 489–515. MR Zbl

[EGA III$_1$ 1961] A. Grothendieck, "Eléments de géométrie algébrique, III: Étude cohomologique des faisceaux cohérents, I", *Inst. Hautes Études Sci. Publ. Math.* **11** (1961), 5–167. MR Zbl

[EGA IV$_1$ 1964] A. Grothendieck, "Eléments de géométrie algébrique, IV: Étude locale des schémas et des morphismes de schémas, I", *Inst. Hautes Études Sci. Publ. Math.* **20** (1964), 5–259. 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

[Ferrand 2003] D. Ferrand, "Conducteur, descente et pincement", *Bull. Soc. Math. France* **131**:4 (2003), 553–585. MR

[Flenner 1981] H. Flenner, "Ein Kriterium für die Offenheit der Versalität", *Math. Z.* **178**:4 (1981), 449–473. MR Zbl

[Hall 2014] J. Hall, "Cohomology and base change for algebraic stacks", *Math. Z.* **278**:1-2 (2014), 401–429. MR Zbl

[Hall 2017] J. Hall, "Openness of versality via coherent functors", *J. Reine Angew. Math.* **722** (2017), 137–182. MR Zbl

[Hall and Rydh 2013] J. Hall and D. Rydh, "Coherence criteria for half-exact functors", article in preparation, latest draft 2013.

[Laumon and Moret-Bailly 2000] G. Laumon and L. Moret-Bailly, *Champs algébriques*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. **39**, Springer, Berlin, 2000. MR Zbl

[Lazard 1964] D. Lazard, "Sur les modules plats", *C. R. Acad. Sci. Paris* **258** (1964), 6313–6316. MR Zbl

[Ogus and Bergman 1972] A. Ogus and G. Bergman, "Nakayama's lemma for half-exact functors", *Proc. Amer. Math. Soc.* **31** (1972), 67–74. MR Zbl

[Olsson 2006] M. C. Olsson, "Deformation theory of representable morphisms of algebraic stacks", *Math. Z.* **253**:1 (2006), 25–62. MR Zbl

[Popescu 1986] D. Popescu, "General Néron desingularization and approximation", *Nagoya Math. J.* **104** (1986), 85–115. MR Zbl

[Schlessinger 1968] M. Schlessinger, "Functors of Artin rings", *Trans. Amer. Math. Soc.* **130** (1968), 208–222. MR Zbl

[SGA 7$_{\mathrm{I}}$ 1972] A. Grothendieck, *Groupes de monodromie en géométrie algébrique, I: Exposés I–II, VI–IX* (Séminaire de Géométrie Algébrique du Bois Marie 1967–1969), Lecture Notes in Math. **288**, Springer, 1972. MR Zbl

[Stacks Project] A. J. de Jong et al., "Stacks Project", electronic reference, http://stacks.math.columbia.edu.

[Starr 2006] J. M. Starr, "Artin's axioms, composition and moduli spaces", 2006. arXiv math/0602646

[Wise 2011] J. Wise, "Obstruction theories and virtual fundamental classes", 2011. arXiv 1111.4200

Communicated by Bjorn Poonen
Received 2013-07-04 Revised 2018-05-28 Accepted 2018-08-03

jackhall@math.arizona.edu *Department of Mathematics, University of Arizona, Tucson, AZ, United States*

dary@math.kth.se *Department of Mathematics, KTH Royal Institute of Technology, Stockholm, Sweden*

msp

ALGEBRA AND NUMBER THEORY 13:4 (2019)
dx.doi.org/10.2140/ant.2019.13.797

Differential characters of Drinfeld modules and de Rham cohomology

James Borger and Arnab Saha

We introduce differential characters of Drinfeld modules. These are function-field analogues of Buium's p-adic differential characters of elliptic curves and of Manin's differential characters of elliptic curves in differential algebra, both of which have had notable Diophantine applications. We determine the structure of the group of differential characters. This shows the existence of a family of interesting differential modular functions on the moduli of Drinfeld modules. It also leads to a canonical F-crystal equipped with a map to the de Rham cohomology of the Drinfeld module. This F-crystal is of a differential-algebraic nature and the relation to the classical cohomological realizations is presently not clear.

1. Introduction

The theory of arithmetic jet spaces developed by Buium draws inspiration from the theory of differential algebra over a function field. In differential algebra, given a scheme E defined over a function field K with a derivation ∂ on it, one can define the jet spaces J^nE for all $n \in \mathbb{N}$ with respect to (K, ∂) and they form an inverse system of schemes satisfying a universal property with respect to derivations lifting ∂. The ring of global functions $\mathcal{O}(J^nE)$ can be thought of as the ring of n-th order differential functions on E. In the case when E is an elliptic curve and its structure sheaf $\mathcal{O}_E$ does not have a derivation lifting ∂ (if it does, then it is the isotrivial case and E will descend to the subfield $K^{\partial=0}$ of constants), there exists a differential function $\Theta \in \mathcal{O}(J^2E)$ which is a homomorphism of group schemes from J^2E to the additive group $\mathbb{G}_a$. Such a Θ is an example of a differential character of order 2 for E and is known as a Manin character. Explicitly, if E is given by the Legendre equation $y^2 = x(x-1)(x-t)$ over $K = \mathbb{C}(t)$ with derivation $\partial = \frac{d}{dt}$, then

$$\Theta(x, y, x', y', x'', y'') = \frac{y}{2(x-t)^2} - \frac{d}{dt}\left[2t(t-1)\frac{x'}{y}\right] + 2t(t-1)x'\frac{y'}{y^2}.$$

The existence of such a Θ is a consequence of the Picard–Fuchs equation. Using the derivation ∂ on K, we can lift any K-rational point $P \in E(K)$ canonically to $J^2E(K)$, and this defines a homomorphism $\nabla : E(K) \to J^2E(K)$. We emphasize that ∇ is merely a map on K-rational points and does not come from a map of schemes. The composition $\Theta \circ \nabla : E(K) \to \mathbb{G}_a(K)$ is then a group homomorphism of

MSC2010: primary 11G99; secondary 14L05.

Keywords: arithmetic geometry, number theory, algebraic geometry, arithmetic jet spaces, Witt vectors, Drinfeld modules, differential characters, de Rham cohomology.

K-points. Note that the torsion points of $E(K)$ are contained in the kernel of Θ since $\mathbb{G}_a(K)$ is torsion free. Such a Θ was used by Manin [1963] to give a proof of the Lang–Mordell conjecture for abelian varieties over function fields. Later Buium [1992] gave a different proof, using other methods, but still using the Manin map.

The theory of arithmetic jet spaces, as developed by Buium, proceeds similarly. Derivations ∂ are replaced by what are known as π-derivations δ. They naturally arise from the theory of π-typical Witt vectors. For instance, when our base ring R is an unramified extension of the ring of p-adic integers $\mathbb{Z}_p$, for a fixed prime $\pi = p$, the Fermat quotient operator $\delta x = (\phi(x) - x^p)/p$ is the unique p-derivation, where the endomorphism $\phi\colon R \to R$ is the lift of the p-th power Frobenius endomorphism of R/pR. In analogy with differential algebra, one can define the n-th order jet space $J^n E$ of an elliptic curve E over R to be the (π-adic) formal scheme over R with functor of points

$$(J^n E)(C) = \mathrm{Hom}_R(\mathrm{Spec}\, W_n(C), E),$$

where $W_n(C)$ is the ring of π-typical Witt vectors of length $n+1$, which we view as the arithmetic analogue of $C[t]/(t^{n+1})$. The jet space $J^n E$ is also known as the Greenberg transform. As with the differential jet space, it has relative dimension $n+1$ over the base, in this case Spf R.

Then one can define $\boldsymbol{X}_n(E)$ to be the R-module of all group-scheme homomorphisms from $J^n E$ to the π-adic formal scheme $\hat{\mathbb{G}}_{\mathrm{a}}$. Let $\boldsymbol{X}_\infty(E)$ be the direct limit of the $\boldsymbol{X}_n(E)$. Now the usual Frobenius operator on Witt vectors induces a canonical Frobenius morphism $\phi : J^{n+1}E \to J^n E$ lying over the endomorphism ϕ of Spf R. Hence pulling back morphisms via ϕ as $\Theta \mapsto \phi^*\Theta$, endows $\boldsymbol{X}_\infty(E)$ with an action of ϕ^* and hence makes $\boldsymbol{X}_\infty(E)$ into a left module over the twisted polynomial ring $R\{\phi^*\}$ with commutation law $\phi^* \cdot r = \phi(r) \cdot \phi^*$. Buium [1995] studied the structure of $\boldsymbol{X}_\infty(E)$. Putting $K = R\big[\frac{1}{p}\big]$, he showed that $\boldsymbol{X}_\infty(E) \otimes_R K$ is freely generated by a single element as a $K\{\phi^*\}$-module. This element is of order 2 unless E has a Frobenius lift (in particular is a canonical lift of an ordinary curve), in which case it is of order 1. It is the arithmetic analogue of the Manin character.

In this paper, we study the function-field analogue of Buium's theory. We emphasize that we take the function-field analogue in every possible sense. So instead of looking at characters $J^n E \to \hat{\mathbb{G}}_{\mathrm{a}}$ of $\mathbb{Z}$-module schemes over $\mathbb{Z}_p$, where the $\mathbb{Z}$-module scheme E is an elliptic curve over $\mathbb{Z}_p$ and $J^n E$ is its p-typical arithmetic jet space defined above, we will look at, for example, characters $J^n E \to \hat{\mathbb{G}}_{\mathrm{a}}$ of (t-adically formal) $\mathbb{F}_q[t]$-module schemes over $\mathbb{F}_q[\![t]\!]$, where E is a Drinfeld $\mathbb{F}_q[t]$-module, $\hat{\mathbb{G}}_{\mathrm{a}}$ is the additive group with the tautological $\mathbb{F}_q[t]$-module structure, and $J^n E$ is its function-field arithmetic jet space — in other words, the Greenberg transform but with "t-typical" Witt vectors. The most important result in this paper is the construction of a canonical F-crystal $\boldsymbol{H}(E)$ which comes with a Hodge-type filtration and a morphism $\boldsymbol{H}(E) \to \boldsymbol{H}_{\mathrm{dR}}(E)$ to the usual de Rham cohomology preserving the filtration. As a consequence of the methods that go into the construction of $\boldsymbol{H}(E)$, we also prove that $\boldsymbol{X}_\infty(E)$ is freely generated by a single element as an $R\{\phi^*\}$-module, which is a stronger, integral version of the equal-characteristic analogue of Buium's result. Here, we would like to emphasize that all the fundamental principles that go into our approach also work for p-adic elliptic curves.

Before we describe our main results in detail, we wish to fix a few notations. Let $\mathbb{F}_q$ be the finite field with q elements and A is the coordinate ring of $X\backslash\{\infty\}$, where X is a projective, geometrically connected, smooth curve over $\mathbb{F}_q$ and ∞ a $\mathbb{F}_q$-point on it. Let $\mathfrak{p}$ be a fixed maximal ideal of A, and let π be an element of $\mathfrak{p}\setminus\mathfrak{p}^2$. Let R be an A-algebra which is a complete discrete valuation ring with maximal ideal πR and which has a lift $\phi : R \to R$ of the $\hat{q}$-power Frobenius from $R/\pi R$, where $\hat{q} = |A/\mathfrak{p}|$. Then one can consider the operator on R given by $\delta x = (\phi(x) - x^{\hat{q}})/\pi$. It is called the *$\pi$-derivation* associated to ϕ.

Then as in the mixed-characteristic case above, one can define the t-typical Witt vectors and hence the t-typical arithmetic jet space functor. For any (formal) A-module scheme E over R, the jet space also J^nE has a natural (formal) A-module-scheme structure. However, we would like to remark here that for all $n \geq 1$, the J^nE are not abelian Anderson A-modules (as defined in [Hartl 2017, 1.2]). Then we let $\boldsymbol{X}_n(E)$ denote the set of A-linear differential characters of order n, that is, the set of homomorphisms $J^nE \to \hat{\mathbb{G}}_{\mathrm{a}}$ of (formal) A-module schemes over R. Finally, we form their direct limit $\boldsymbol{X}_\infty(E)$, which is naturally an $R\{\phi^*\}$-module, as above.

We say E splits at m if $\boldsymbol{X}_m(E) \neq \{0\}$ but $\boldsymbol{X}_i(E) = \{0\}$ for all $0 \leq i \leq m-1$. Then we show that m satisfies $1 \leq m \leq r$, where r is the rank of E, and that $\boldsymbol{X}_m(E)$ is a free R-module with a canonical basis element $\Theta_m \in \boldsymbol{X}_m(E)$, depending only on our chosen coordinate on E. In the case when the rank r is 2, we have $m = 2$ unless E admits a lift of Frobenius compatible with the A-module structure on E, in which case $m = 1$. Then our first main theorem is a strengthened version of the equal-characteristic analogue of Buium's result [1995].

Theorem 1.1. *Let E be a Drinfeld module that splits at m. Then the R-module $\boldsymbol{X}_m(E)$ is free of rank 1 and it freely generates $\boldsymbol{X}_\infty(E)$ as an $R\{\phi^*\}$-module in the sense that the canonical map $R\{\phi^*\} \otimes_R \boldsymbol{X}_m(E) \to \boldsymbol{X}_\infty(E)$ is an isomorphism.*

Let us now proceed to our second result. Let $u : J^nE \to E$ be the usual projection map and put $N^n = \ker u$. Since u is A-linear, N^n is a formal A-module scheme of relative dimension n over $\operatorname{Spf} R$. For each $n \geq 1$, we show in Proposition 7.2 that there is a lift of Frobenius $\mathfrak{f} : N^{n+1} \to N^n$ making the system $\{N^n\}$ into a prolongation sequence with respect the obvious projection map $u : N^{n+1} \to N^n$. We call $\mathfrak{f}$ the *lateral Frobenius*. However, $\mathfrak{f}$ is not compatible with i and $\phi : J^{n+1}E \to J^nE$ in the obvious way, that is, it is not true that $\phi \circ i = i \circ \mathfrak{f}$ holds. In fact, we can not expect it to be true because that would induce an A-linear lift of Frobenius on E which is not the case to start with. Instead we have

$$\phi^2 \circ i = \phi \circ i \circ \mathfrak{f}.$$

In Section 9, we construct a canonical F-crystal attached to E. The F-crystal, denoted $\boldsymbol{H}(E)$, is an R-module which has a semilinear operator $\mathfrak{f}^*$ (induced from $\mathfrak{f}$) on it and is of rank m, which we emphasize can be strictly smaller than r. (By the term F-crystal, we mean only a free R-module of finite rank equipped with a semilinear operator F. We do not assume F is injective, although on $\boldsymbol{H}(E)$ this will be true generically. The reader can refer to [Laumon 1996, §2.4].) The module $\boldsymbol{H}(E)$ also has a Hodge-type filtration and canonically maps to the de Rham cohomology of E, with its Hodge filtration.

Theorem 1.2. *There is a canonical map between exact sequences*

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \boldsymbol{X}_m(E) & \longrightarrow & \boldsymbol{H}(E) & \longrightarrow & \boldsymbol{I}(E) & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle \Upsilon} & & \downarrow{\scriptstyle \Phi} & & \downarrow & & \\ 0 & \longrightarrow & \mathrm{Lie}(E)^* & \longrightarrow & \boldsymbol{H}_{\mathrm{dR}}(E) & \longrightarrow & \mathrm{Ext}(E, \hat{\mathbb{G}}_{\mathrm{a}}) & \longrightarrow & 0 \end{array}$$

Moreover, the operator $\mathfrak{f}^*$ *on* $\boldsymbol{H}(E)$ *descends to its image under* Φ.

The definitions of the maps Υ and Φ are given in (9-7), and the proof is given in Section 9B. There is a close connection between these two theorems — in fact, our proof of Theorem 1.1 goes by way of Theorem 1.2.

Finally, we conclude the paper with some explicit computations of the structure constants of the F-crystal $\boldsymbol{H}(E)$, which are new differential modular forms.

To a Drinfeld module E, the crystalline theory also attaches an F-crystal $\boldsymbol{H}_{\mathrm{crys}}(E)$. It appears that our $\boldsymbol{H}(E)$ has subtle connections with $\boldsymbol{H}_{\mathrm{crys}}(E)$, but it also appears that any such connection would be indirect. This is because $\boldsymbol{H}(E)$, unlike $\boldsymbol{H}_{\mathrm{crys}}(E)$, has a fundamentally differential-algebraic nature in that it lies not over a point of the moduli space of Drinfeld modules but over a point of the jet space of the moduli space. For instance, the computations in Section 10 show the structure constants of $\boldsymbol{H}(E)$ do involve the higher π-derivatives of the structure constants of the Drinfeld module. The phenomenon of π-differential invariants depending on higher π-derivatives of modular parameters in the mixed-characteristic setting can be found in [Borger and Saha 2017a; Buium 1995; Buium and Saha 2011; 2012a; 2012b; 2014].

It would be interesting to understand the exact nature of the relationship between $\boldsymbol{H}(E)$ and the crystalline cohomology groups, as well as the étale cohomology groups and the other constructions in π-adic Hodge theory. This is all the more true because, as we remarked before, the techniques developed in this paper have analogues for p-adic elliptic curves [Borger and Saha 2017a], and as a result, we do obtain an analogous construction of the F-crystal $\boldsymbol{H}(E)$ for elliptic curves.

2. Notation

Let us fix some notation which will hold throughout the paper. Let $q = p^h$ where p is a prime and $h \geq 1$. Let X be a projective, geometrically connected, smooth curve over $\mathbb{F}_q$. Fix an $\mathbb{F}_q$-rational point ∞ on X. Let A denote the Dedekind domain $\mathcal{O}(X \setminus \{\infty\})$. Let $\mathfrak{p}$ be a maximal ideal of A, and let $\hat{A}$ denote the $\mathfrak{p}$-adic completion of A. Let t be an element of $\mathfrak{p} \setminus \mathfrak{p}^2$, and let π denote its image in $\hat{A}$. Then π generates the maximal ideal $\hat{\mathfrak{p}}$ of $\hat{A}$. Let k denote the residue field $A/\mathfrak{p}$ and let $\hat{q}$ denote its cardinality. So, for example, if $A = \mathbb{F}_q[u]$ and $\mathfrak{p} = (t)$, where $t \in \mathbb{F}_q[u]$ is an irreducible polynomial, then $\hat{q} = q^{\deg(t)}$. Note that the quotient map $\hat{A} \to k$ has a unique section. Thus $\hat{A}$ is not just an $\mathbb{F}_q$-algebra but also canonically a k-algebra.

Now let R be an $\hat{A}$-algebra which is $\mathfrak{p}$-adically complete and flat, or equivalently π-torsion free. Thus the composition

$$\theta : A \to \hat{A} \to R \tag{2-1}$$

is injective (assuming $R \neq \{0\}$) and hence one says that θ is of *generic characteristic*. Let us also fix an $\hat{A}$-algebra endomorphism $\phi : R \to R$ which lifts the $\hat{q}$-power Frobenius modulo $\mathfrak{p}R$:

$$\phi(x) \equiv x^{\hat{q}} \bmod \mathfrak{p}R.$$

Do note that the identity map on $\hat{A}$ does indeed lift the $\hat{q}$-power Frobenius on $\hat{A}/\hat{\mathfrak{p}}$.

For our main results, R will in the end be a discrete valuation ring, most importantly the completion $\bar{\mathbb{F}}_q[\![\pi]\!]$ of the maximal unramified extension of $\hat{A}$, where ϕ satisfies $\phi(c) = c^{\hat{q}}$ for $c \in \bar{\mathbb{F}}_q$ and $\phi(\pi) = \pi$. So the reader may assume this from the start. (Also note that not all rings R admit such a Frobenius lift; so the existence of ϕ does place a restriction on R.) But some form of our results should hold in general, and with essentially the same proofs. This is of some interest, for instance when R is the coordinate ring of the ordinary locus of the moduli space of Drinfeld modules of a given rank. (For the representability of Drinfeld modular varieties, see Laumon's book [1996, Theorem 1.4.1].) With an eye to the future, we have not assumed that R is a discrete valuation ring where it is easily avoided, in Sections 3–7.

Let K denote $R\big[frac1\pi\big]$, and for any R-module M write $M_K = K \otimes_R M$. Finally, let S denote Spf R.

3. Function-field Witt vectors

Witt vectors over Dedekind domains with finite residue fields were introduced in [Borger 2011a]. We will only work over $\hat{A}$, which is the ring of integers of a local field of characteristic p, and here they were introduced earlier in [Drinfeld 1976]. The basic results can be developed exactly as in any of the usual developments of the p-typical Witt vectors. The only difference is that in all formulas any p in a coefficient is replaced with a π and any p in an exponent is replaced with a $\hat{q}$.

3A. ***Frobenius lifts and π-derivations.*** Let B be an R-algebra, and let C be a B-algebra with structure map $u : B \to C$. In this paper, a ring homomorphism $\psi : B \to C$ will be called a *lift of Frobenius* (relative to u) if it satisfies the following:

(1) The reduction mod π of ψ is the $\hat{q}$-power Frobenius relative to u, that is, $\psi(x) \equiv u(x)^{\hat{q}} \bmod \pi C$.

(2) The restriction of ψ to R coincides with the fixed ϕ on R, that is, the following diagram commutes

$$\begin{array}{ccc} B & \xrightarrow{\psi} & C \\ \uparrow & & \uparrow \\ R & \xrightarrow[\phi]{} & R \end{array}$$

A *π-derivation* δ from B to C means a set-theoretic map $\delta : B \to C$ satisfying the following for all $x, y \in B$

$$\delta(x+y) = \delta(x) + \delta(y) \quad \text{and} \quad \delta(xy) = u(x)^{\hat{q}}\delta(y) + \delta(x)u(y)^{\hat{q}} + \pi\delta(x)\delta(y)$$

such that for all $r \in R$, we have

$$\delta(r) = \frac{\phi(r) - r^{\hat{q}}}{\pi}.$$

When $C = B$ and u is the identity map, we will call this simply a π-derivation on B.

It follows that the map $\phi : B \to C$ defined as

$$\phi(x) := u(x)^{\hat{q}} + \pi\delta(x)$$

is a lift of Frobenius in the sense above. On the other hand, for any flat R-algebra B with a lift of Frobenius ϕ, one can define the π-derivation $\delta(x) = (\phi(x) - x^{\hat{q}})/\pi$ for all $x \in B$.

Note that this definition depends on the choice of uniformizer π, but in a transparent way: if π' is another uniformizer, then $\delta(x)\pi/\pi'$ is a π'-derivation. This correspondence induces a bijection between π-derivations $B \to C$ and π'-derivations $B \to C$.

3B. ***Witt vectors.*** We will present three different points of view on function-field Witt vectors, all parallel to the mixed characteristic case. But there is perhaps one unfamiliar element below, which is that we will work relative to our general base R, and it already has a lift of Frobenius. The consequence is that we need to pay attention to certain twists of the scalars by Frobenius, which are invisible over the absolute base $R = \hat{A}$. However this unfamiliar element has nothing to do with the difference between mixed and equal characteristic and only with the difference between the relative and the absolute setting.

Let B be an R-algebra with structure map $u : R \to B$.

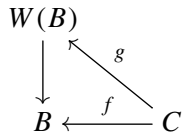
(1) The ring $W(B)$ of *π-typical Witt vectors* can be defined as the unique (up to unique isomorphism) R-algebra $W(B)$ with a π-derivation δ on $W(B)$ and an R-algebra homomorphism $W(B) \to B$ such that, given any R-algebra C with a π-derivation δ on it and an R-algebra map $f : C \to B$, there exists a unique R-algebra homomorphism $g : C \to W(B)$ such that the diagram

$$\begin{array}{ccc} W(B) & & \\ \downarrow & \nwarrow^{g} & \\ B & \xleftarrow{f} & C \end{array}$$

commutes and $g \circ \delta = \delta \circ g$. Thus W is the right adjoint of the forgetful functor from R-algebras with π-derivation to R-algebras. For details, see Section 1 of [Borger 2011a]. This approach follows that of [Joyal 1985] to the usual p-typical Witt vectors.

(2) If we restrict to flat R-algebras B, then we can ignore the concept of π-derivation and define $W(B)$ simply by expressing the universal property above in terms of Frobenius lifts, as follows. Given a flat R-algebra B, the ring $W(B)$ is the unique (up to unique isomorphism) flat R-algebra $W(B)$ with a lift of Frobenius (in the sense above) $F : W(B) \to W(B)$ and an R-algebra homomorphism $W(B) \to B$ such that for any flat R-algebra C with a lift of Frobenius ϕ on it and an R-algebra map $f : C \to B$, there

exists a unique R-algebra homomorphism $g : C \to W(B)$ such that the diagram



commutes and $g \circ \phi = F \circ g$.

(3) Finally, returning to the case of general R-algebras B, one can also define Witt vectors in terms of the Witt polynomials. For each $n \geq 0$ let us define B^{ϕ^n} to be the R-algebra with structure map $R \xrightarrow{\phi^n} R \xrightarrow{u} B$ and define the *ghost rings* to be the product R-algebras $\prod_\phi^n B = B \times B^\phi \times \cdots \times B^{\phi^n}$ and $\prod_\phi^\infty B = B \times B^\phi \times \cdots$. Then for all $n \geq 1$ there exists a *restriction*, or *truncation*, map $T_w : \prod_\phi^n B \to \prod_\phi^{n-1} B$ given by $T_w(w_0, \cdots, w_n) = (w_0, \cdots, w_{n-1})$. We also have the left shift *Frobenius* operators $F_w : \prod_\phi^n B \to \prod_\phi^{n-1} B$ given by $F_w(w_0, \ldots, w_n) = (w_1, \ldots, w_n)$. Note that T_w is an R-algebra morphism, but F_w lies over the Frobenius endomorphism ϕ of R.

Now as sets define

$$W_n(B) = B^{n+1}, \tag{3-1}$$

and define the set map $w : W_n(B) \to \prod_\phi^n B$ by $w(x_0, \ldots, x_n) = (w_0, \ldots, w_n)$ where

$$w_i = x_0^{\hat{q}^i} + \pi x_1^{\hat{q}^{i-1}} + \cdots + \pi^i x_i \tag{3-2}$$

are the *Witt polynomials*. The map w is known as the *ghost* map. (Do note that under the traditional indexing, used in many sources going back to Witt [1937], our W_n would be denoted W_{n+1}.) We can then define the ring $W_n(B)$, the ring of truncated π-typical Witt vectors, by the following theorem as in the p-typical case [Hesselholt 2015, Proposition 1.2].

Theorem 3.1. *For each $n \geq 0$, there exists a unique functorial R-algebra structure on $W_n(B)$ such that w becomes a natural transformation of functors of R-algebras.*

Note that, unlike with Witt vectors in mixed characteristic, addition for function-field Witt vectors is performed componentwise. This is because the Witt polynomials (3-2) are additive. This might appear to defeat the whole point of Witt vectors and arithmetic jet spaces. But this is not so. The reason is that while the additive structure is the componentwise one, the A-module structure is not. So the difference is only that, unlike in mixed characteristic where $A = \mathbb{Z}$, a group structure is weaker than A-module structure. In fact, because the Witt polynomials are k-linear, the k-vector space structure on $W_n(B)$ is the componentwise one. This is just like with the p-typical Witt vectors, where multiplication by roots of $x^p - x$ can be performed componentwise.

For the convenience of the reader, we give some examples the proofs of which we leave as exercises. If the structure map $A \to B$ factors through $A/\mathfrak{p}$ and B is perfect, then multiplication is given by the formula

$$(x_0, x_1, \ldots) \cdot (y_0, y_1, \ldots) = (z_0, z_1, \ldots), \quad \text{where } z_n = \sum_{i+j=n} x_i^{\hat{q}^j} y_j^{\hat{q}^i}.$$

For example, if $B = R = A/\mathfrak{p} = \mathbb{F}_{\hat{q}}$, then $W(B)$ is identified with the power-series ring $B[\![\pi]\!]$, where π corresponds to the Witt vector $(0, 1, 0, 0, \ldots)$. At the opposite extreme, where π is invertible in B, the ghost map is an isomorphism. So $W(B)$ is isomorphic to the product ring $B \times B \times \cdots$ and not a power-series ring.

3C. ***Operations on Witt vectors.*** Now we recall some important operators on the Witt vectors. There are the *restriction*, or *truncation*, maps $T : W_n(B) \to W_{n-1}(B)$ given by $T(x_0, \ldots, x_n) = (x_0, \ldots, x_{n-1})$. Note that $W(B) = \varprojlim W_n(B)$. There is also the *Frobenius* ring homomorphism $F : W_n(B) \to W_{n-1}(B)$, which can be described in terms of the ghost map. It is the unique map which is functorial in B and makes the following diagram commutative

$$\begin{array}{ccc} W_n(B) & \xrightarrow{w} & \prod_\phi^n B \\ {\scriptstyle F}\downarrow & & \downarrow{\scriptstyle F_w} \\ W_{n-1}(B) & \xrightarrow[w]{} & \prod_\phi^{n-1} B^n \end{array} \tag{3-3}$$

As with the ghost components, T is an R-algebra map but F lies over the Frobenius endomorphism ϕ of R.

Next we have the *Verschiebung* $V : W_{n-1}(B) \to W_n(B)$ given by

$$V(x_0, \ldots, x_{n-1}) = (0, x_0, \ldots, x_{n-1}).$$

Let $V_w : \prod_\phi^{n-1} B \to \prod_\phi^n B$ be the additive map given by

$$V_w(w_0, \ldots, w_{n-1}) = (0, \pi w_0, \ldots, \pi w_{n-1}).$$

Then the Verschiebung V makes the following diagram commute:

$$\begin{array}{ccc} W_{n-1}(B) & \xrightarrow{w} & \prod_\phi^{n-1} B \\ {\scriptstyle V}\downarrow & & \downarrow{\scriptstyle V_w} \\ W_n(B) & \xrightarrow[w]{} & \prod_\phi^n B \end{array} \tag{3-4}$$

For all $n \geq 0$ the Frobenius and the Verschiebung satisfy the identity

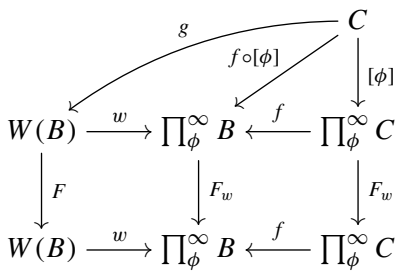
$$FV(x) = \pi x. \tag{3-5}$$

The Verschiebung is not a ring homomorphism, but it is k-linear.

Finally, we have the multiplicative Teichmüller map $[\ ] : B \to W_n(B)$ given by $x \mapsto [x] = (x, 0, 0, \ldots)$. Here in the function-field setting, $[\ ]$ is additive and even a homomorphism of k-algebras but is not a homomorphism of A-algebras. This can be compared to the mixed-characteristic setting, where it is a homomorphism of monoids but not a homomorphism of $\mathbb{Z}$-algebras.

3D. ***Computing the universal map to Witt vectors.*** Given an R-algebra C with a π-derivation $\delta : C \to C$ and an R-algebra map $f : C \to B$, we will now describe the universal lift $g : C \to W(B)$. The explicit description of g leads us to Proposition 3.2 which is used in Section 10 in computations for Drinfeld modules of rank 2. The reader may skip this subsection without breaking continuity till then.

It is enough to work in the case where both B and C are flat over R. Then the ghost map $w : W(B) \to \prod_\phi^\infty B$ is injective. Consider the map $[\phi] : C \to \prod_\phi^\infty C$ given by $x \mapsto (x, \phi(x), \phi^2(x), \ldots)$. Then we have the following commutative diagram:



Thus the map $f \circ [\phi] : C \to \prod_\phi^\infty B$ factors through $W(B)$ as our universal map $g : C \to W(B)$.

Let us now give an inductive description of the map g. Write

$$g(x) = (x_0, x_1, \ldots) \in W(B).$$

Then from the above diagram $w \circ g = f \circ [\phi]$. Therefore the vector $(x_0, x_1, \ldots)$ is the unique solution to the system of equations

$$x_0^{\hat{q}^n} + \pi x_1^{\hat{q}^{n-1}} + \cdots + \pi^n x_n = f(\phi^n(x)), \tag{3-6}$$

for $n \geq 0$. For example, we have $x_0 = f(x)$ and $x_1 = f(\delta(x))$.

Now consider the case where B itself has a π-derivation, $C = B$, and $f = \mathbb{1}$. For any $x \in B$, let us write $x^{(n)} := \delta^n(x)$, or simply $x' = \delta(x)$, $x'' = \delta^2(x)$ and so on.

Proposition 3.2. *We have $x_0 = x$, $x_1 = x'$ and $x_2 = x'' + \pi^{\hat{q}-2}(x')^{\hat{q}}$.*

Proof. As stated above, equalities $x_0 = x$ and $x_1 = x'$ follow immediately from (3-6). For $n = 2$, we have

$$\begin{aligned} x_0^{\hat{q}^2} + \pi x_1^{\hat{q}} + \pi^2 x_2 &= \phi^2(x) \\ &= \phi(x^{\hat{q}} + \pi x') \\ &= \phi(x)^{\hat{q}} + \pi\phi(x') \\ &= x^{\hat{q}^2} + \pi^{\hat{q}}(x')^{\hat{q}} + \pi((x')^{\hat{q}} + \pi x'') \end{aligned}$$

And therefore we have $x_2 = x'' + \pi^{\hat{q}-2}(x')^{\hat{q}}$. □

4. A-module schemes, jet spaces and preliminaries

An A-module scheme over $S = \operatorname{Spf} R$ is by definition a pair (E, φ_E), where E is a commutative group object in the category of S-schemes and $\varphi_E : A \to \operatorname{End}(E/S)$ is a ring map. (Here and below, by a scheme over the formal scheme S, we mean a formal scheme formed from a compatible family of schemes over the schemes $\operatorname{Spec} R/\mathfrak{p}^n R$.) Then the tangent space T_0E at the identity has two A-modules structures: one coming by restriction of the usual R-module structure to A, and the other coming from differentiating φ_E. We will say that (E, φ_E) is *strict* if these two A-module structures coincide, that is if the composition

$$A \to \operatorname{End}(E/S) \to \operatorname{End}_R(T_0E)$$

agrees with the composition

$$A \xrightarrow{\theta} R \to \operatorname{End}_R(T_0E).$$

We say it is *admissible* if it is both strict and isomorphic to the additive group $\hat{\mathbb{G}}_{\mathrm{a}} = \hat{\mathbb{G}}_{\mathrm{a}/S}$ as a group scheme.

We will denote this induced map to tangent space as $\theta : A \to R$. (Note that it is best practice to require only the isomorphism with $\hat{\mathbb{G}}_{\mathrm{a}}$ to exist locally on S. So below, our Drinfeld modules would more properly be called *coordinatized* Drinfeld modules.)

A *Drinfeld module* (E, φ_E) *of rank* r is an admissible A-module scheme over S such that for each nonzero $a \in A$, the group scheme $\ker(\varphi_E(a))$ is finite flat of degree $|a|^r = q^{-r\operatorname{ord}_\infty(a)}$ over S. (See [Gekeler 1990b, (1.4)] or [Laumon 1996, p. 4].)

Proposition 4.1. *Let f be an endomorphism of the $\mathbb{F}_q$-module scheme $\hat{\mathbb{G}}_{\mathrm{a}/S}$ over S. Then given any coordinate x on E, the map f is of the form*

$$f(x) = \sum_{i=0}^{\infty} a_i x^{q^i},$$

where f is a restricted power series, meaning $a_i \to 0$ π-adically as $i \to \infty$.

Proof. Let $f \in \operatorname{Hom}(\hat{\mathbb{G}}_{\mathrm{a}}, \hat{\mathbb{G}}_{\mathrm{a}})$ be an additive endomorphism of $\hat{\mathbb{G}}_{\mathrm{a}}$. Then f is given a restricted power series $\sum_i b_i x^i$ such that $b_i \to 0$ as $i \to \infty$. Since f is additive, we have $b_i = 0$ unless i is a power of p. Second, because f is $\mathbb{F}_q$-linear, we have $\sum_i b_{p^i}(cx)^{p^i} = c\sum_i b_{p^i} x^{p^i}$ for all $c \in \mathbb{F}_q$. Considering the case where c is a generator of $\mathbb{F}_q^*$, we see this implies $b_{p^i} = 0$ unless p^i is a power of q. □

Let $R\{\tau\}\hat{}$ be the subring of $R\{\{\tau\}\}$ consisting of (twisted) restricted power series. Then by Proposition 4.1, the $\mathbb{F}_q$-linear morphisms between two admissible A-module schemes E_1 and E_2 over $\operatorname{Spf} R$ are given in coordinates by elements in $R\{\tau\}\hat{}$ where τ acts as $\tau(x) = x^q$:

$$\operatorname{Hom}_{\mathbb{F}_q}(E_1, E_2) = R\{\tau\}\hat{}. \tag{4-1}$$

4A. ***Prolongation sequences and jet spaces.*** Let X and Y be schemes over $S = \operatorname{Spf} R$. We say a pair (u, δ) is a *prolongation*, and write $Y \xrightarrow{(u,\delta)} X$, if $u : Y \to X$ is a map of schemes over S and $\delta : \mathcal{O}_X \to u_*\mathcal{O}_Y$

is a π-derivation making the following diagram commute:

$$\begin{array}{ccc} R & \longrightarrow & u_*\mathcal{O}_Y \\ {\scriptstyle\delta}\uparrow & & \uparrow{\scriptstyle\delta} \\ R & \longrightarrow & \mathcal{O}_X \end{array}$$

Following [Buium 2000], a *prolongation sequence* is a sequence of prolongations

$$\operatorname{Spf} R \xleftarrow{(u,\delta)} T^0 \xleftarrow{(u,\delta)} T^1 \xleftarrow{(u,\delta)} \cdots,$$

where each T^n is a scheme over S. We will often use the notation T^* or $\{T_n\}_{n\geq 0}$. Note that if the T^n are flat over $\operatorname{Spf} R$ then having a π-derivation δ is equivalent to having lifts of Frobenius $\phi : T^{n+1} \to T^n$.

Prolongation sequences form a category $\mathcal{C}_{S^*}$, where a morphism $f : T^* \to U^*$ is a family of morphisms $f^n : T^n \to U^n$ commuting with both the u and δ, in the evident sense. This category has a final object S^* given by $S^n = \operatorname{Spf} R$ for all n, where each u is the identity and each δ is the given π-derivation on R.

For any scheme Y over S, for all $n \geq 0$ we define the n-th jet space $J^n X$ (relative to S) as

$$J^n X(Y) := \operatorname{Hom}_S(W_n^*(Y), X)$$

where $W_n^*(Y)$ is defined in Section 10.3 of [Borger 2011b]. We will not define $W_n^*(Y)$ in full generality here. Instead, we will define $\operatorname{Hom}_S(W_n^*(Y), X)$ in the affine case, and that will be sufficient for the purposes of this paper. Write $X = \operatorname{Spf} C$ and $Y = \operatorname{Spf} B$. Then $W_n^*(Y) = \operatorname{Spf} W_n(B)$ and so $J^n X(B)$ is the set of R-algebra homomorphisms $C \to W_n(B)$:

$$J^n X(B) = \operatorname{Hom}_R(C, W_n(B)). \tag{4-2}$$

Then $J^* X := \{J^n X\}_{n\geq 0}$ forms a prolongation sequence, called the *canonical prolongation sequence*. As in the mixed-characteristic case [Buium 2000, Proposition 1.1], $J^* X$ satisfies the following universal property — for any $T^* \in \mathcal{C}_{S^*}$ and X a scheme over S^0, we have

$$\operatorname{Hom}(T^0, X) = \operatorname{Hom}_{\mathcal{C}_{S^*}}(T^*, J^* X)$$

Let X be a scheme over $S = \operatorname{Spf} R$. Define X^{ϕ^n} by $X^{\phi^n}(B) := X(B^{\phi^n})$ for any R-algebra B. In other words, X^{ϕ^n} is $X \times_{S,\phi^n} S$, the pull-back of X under the map $\phi^n : S \to S$. Next define

$$\prod_{\phi}^{n} X = X \times_S X^{\phi} \times_S \cdots \times_S X^{\phi^n}.$$

Then for any R-algebra B we have $X\left(\prod_{\phi}^{n} B\right) = X(B) \times_S \cdots \times_S X^{\phi^n}(B)$. Thus the ghost map w in Theorem 3.1 defines a map of S-schemes

$$w : J^n X \to \prod_{\phi}^{n} X.$$

Note that w is injective when evaluated on points with coordinates in any flat R-algebra.

The operators F and F_w in (3-3) induce maps ϕ and ϕ_w as follows

$$\begin{array}{ccc} J^n X & \xrightarrow{w} & \prod_\phi^n X \\ {\scriptstyle\phi}\downarrow & & \downarrow{\scriptstyle\phi_w} \\ J^{n-1}X & \xrightarrow[w]{} & \prod_\phi^{n-1} X \end{array} \tag{4-3}$$

where ϕ_w is the left-shift operator given by

$$\phi_w(w_0, \dots, w_n) = (\phi_S(w_1), \dots, \phi_S(w_n)),$$

and where $\phi_S : X^{\phi^i} \to X^{\phi^{i-1}}$ is the composition given in the following diagram:

$$\begin{array}{ccccc} X^{\phi^i} & \xrightarrow{\sim} & X^{\phi^{i-1}} \times_{S,\phi} S & \longrightarrow & X^{\phi^{i-1}} \\ & & \downarrow & & \downarrow \\ & & S & \xrightarrow[\phi]{} & S. \end{array} \tag{4-4}$$

Now let E be an A-module scheme over S with action map $A \xrightarrow{\varphi_E} \mathrm{End}_S(E)$. Then the functor it represents takes values in A-modules, and hence so does the functor $B \mapsto E(W_n(B))$. In this way, for each $n \geq 0$, the S-scheme $J^n E$ comes with an A-module structure. We denote it by $\varphi_{J^n E} : A \to \mathrm{End}_S(J^n E)$. Similarly, φ_E induces an A-linear structure $\varphi_{E^{\phi^n}}$ on each E^{ϕ^n}. In this case, it is easy to describe explicitly. It is the componentwise one:

$$\varphi_{\prod_\phi^n E}(w_0, \dots, w_n) = (\varphi_E(w_0), \dots, \varphi_{E^{\phi^n}}(w_n)).$$

The ghost map $w : J^n E \to \prod_\phi^n E$ and the truncation map $u : J^n E \to J^{n-1}E$ homomorphisms of A-module schemes over S. This is because they are given by applying the A-module scheme E to the R-algebra maps $w : W_n(B) \to \prod_\phi^n B$ and $T : W_n(B) \to W_{n-1}(B)$. On the other hand, the Frobenius map $\phi : J^n E \to J^{n-1}E$ is a homomorphisms of A-module schemes lying over the Frobenius endomorphism ϕ of S. In other words, the induced map $J^n E \to (J^{n-1}E)^\phi$ is a homomorphism of A-module schemes over S.

4B. ***Coordinates on jet spaces.*** Given an isomorphism of S-schemes $E \to \hat{\mathbb{G}}_{\mathrm{a}}$, we have an induced bijection, by (4-2),

$$(J^n E)(B) \xrightarrow{\sim} W_n(B). \tag{4-5}$$

Now recall the bijection $W_n(B) \xrightarrow{\sim} B^{n+1}$ of (3-1). Combining the two, we see that given a coordinate x on an admissible A-module scheme E, we have a canonical system of coordinates $(x_0, \dots, x_n)$ on $J^n E$. We will use these *Witt coordinates* without further comment. We emphasize once again that there are other canonical systems of coordinates on $J^n E$, for instance the *Buium–Joyal* coordinates denoted $x, x', x'', \dots$. They are related by the formulas of Proposition 3.2. Each has their own advantages.

We will now describe the above maps explicitly in the Buium–Joyal coordinates. Let $\mathcal{O}(E) = R[x]\hat{}$. Then, for each n, $\mathcal{O}(J^n E) = R[x, x', \dots, x^{(n)}]$ and the corresponding algebra maps u^* and ϕ^* from $\mathcal{O}(J^n E) \to \mathcal{O}(J^{n+1} E)$ are given as follows, for all i:

$$\begin{aligned} u^*(x^{(i)}) &= x^{(i)}, \\ \phi^*(x^{(i)}) &= (x^{(i)})^{\hat{q}} + \pi x^{(i+1)}. \end{aligned} \tag{4-6}$$

4C. ***Character groups.*** Let $\hat{\mathbb{G}}_{\mathrm{a}}$ denote the additive group over S, i.e., the formal spectrum of the π-adic completion of $R[x]$, with the tautological A-module structure $\varphi_{\hat{\mathbb{G}}_{\mathrm{a}}}$ given by the usual multiplication of scalars: $\varphi_{\hat{\mathbb{G}}_{\mathrm{a}}}(a) = a\tau^0$. We will maintain this convention throughout the paper.

Given a prolongation sequence T^* we can define its shift T^{*+n} by $(T^{*+n})^j := T^{n+j}$ for all j (as in [Buium 2000, p. 106]).

$$\operatorname{Spf} R \xleftarrow{(u,\delta)} T^n \xleftarrow{(u,\delta)} T^{n+1} \cdots$$

We define a *δ-morphism of order n* from X to Y to be a morphism $J^{*+n}X \to J^*Y$ of prolongation sequences. We define a *character of order n*, $\Theta : (E, \varphi_E) \to (\hat{\mathbb{G}}_{\mathrm{a}}, \varphi_{\hat{\mathbb{G}}_{\mathrm{a}}})$ to be a δ-morphism of order n from E to $\hat{\mathbb{G}}_{\mathrm{a}}$ which is also a homomorphism of A-module objects. By the same argument as in the mixed characteristic case [Buium 2000, Proposition 1.9], an order n character is equivalent to a homomorphism $\Theta : J^n E \to \hat{\mathbb{G}}_{\mathrm{a}}$ of A-module schemes over S. We denote the group of characters of order n by $\boldsymbol{X}_n(E)$. So we have

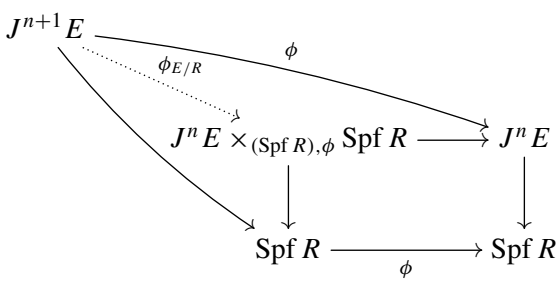
$$\boldsymbol{X}_n(E) = \operatorname{Hom}_A(J^n E, \hat{\mathbb{G}}_{\mathrm{a}}),$$

which one could take as an alternative definition. Note that $\boldsymbol{X}_n(E)$ comes with an R-module structure since $\hat{\mathbb{G}}_{\mathrm{a}}$ is an R-module scheme over S. Also the inverse system $J^{n+1}E \xrightarrow{u} J^n E$ defines a directed system

$$\boldsymbol{X}_n(E) \xrightarrow{u^*} \boldsymbol{X}_{n+1}(E) \xrightarrow{u^*} \cdots$$

via pull back. Each morphism u^* is injective because each u has a section (typically not A-linear). We then define $\boldsymbol{X}_\infty(E)$ to be the R-module direct limit $\varinjlim \boldsymbol{X}_n(E)$.

Similarly, precomposing with the Frobenius map $\phi : J^{n+1}E \to J^n E$ induces a Frobenius operator $\phi : \boldsymbol{X}^n(E) \to \boldsymbol{X}^{n+1}(E)$. However since $\phi : J^{n+1}E \to J^n E$ is not a morphism over $\operatorname{Spf} R$ but instead lies over the Frobenius endomorphism ϕ of $\operatorname{Spf} R$, some care is required. Consider the relative Frobenius morphism $\phi_{E/R}$, defined to be the unique morphism making the following diagram commute:



Then $\phi_{E/R}$ is a morphism of A-module formal schemes over $\operatorname{Spf} R$. Now given a δ-character $\Theta : J^n E \to \hat{\mathbb{G}}_a$, define $\phi^*\Theta$ to be the composition

$$J^{n+1}E \xrightarrow{\phi_{E/R}} J^n E \times_{(\operatorname{Spf} R),\phi} \operatorname{Spf} R \xrightarrow{\Theta\times\mathbb{1}} \hat{\mathbb{G}}_a \times_{(\operatorname{Spf} R),\phi} \operatorname{Spf} R \xrightarrow{\iota} \hat{\mathbb{G}}_a, \tag{4-7}$$

where ι is the isomorphism of A-module schemes over S coming from the fact that $\hat{\mathbb{G}}_a$ descends to $\hat{A}$ as an A-module scheme. For any R-algebra B, the induced morphism on B-points is

$$E(W_{n+1}(B)) \xrightarrow{E(F)} E(W_n(B)^\phi) \xrightarrow{\Theta_B^\phi} B^\phi \xrightarrow{b\mapsto b} B.$$

Note that this composition $E(W_{n+1}(B)) \to B$ is indeed a morphism of A-modules because identity map $B^\phi \to B$ is A-linear, which is true because ϕ restricted to $\hat{A}$ is the identity.

Thus we have an additive map $\boldsymbol{X}_n(E) \to \boldsymbol{X}_{n+1}(E)$ given by $\Theta \mapsto \phi^*\Theta$. Note that this map is not R-linear. However, the map

$$\phi^* : \boldsymbol{X}_n(E) \to \boldsymbol{X}_{n+1}(E)^\phi, \quad \Theta \mapsto \phi^*\Theta$$

is R-linear, where $\boldsymbol{X}_{n+1}(E)^\phi$ denotes the abelian group $\boldsymbol{X}_{n+1}(E)$ with R-module structure defined by the law $r \cdot \Theta := \phi(r)\Theta$. Taking direct limits in n, we obtain an R-linear map

$$\boldsymbol{X}_\infty(E) \to \boldsymbol{X}_\infty(E)^\phi, \quad \Theta \mapsto \phi^*\Theta.$$

In this way, $\boldsymbol{X}_\infty(E)$ is a left module over the twisted polynomial ring $R\{\phi^*\}$ with commutation law $\phi^* r = \phi(r)\phi^*$.

5. Admissible modules

Let (E, φ_E) be an admissible A-module scheme over $S = \operatorname{Spf} R$. By (4-1), we can write

$$\varphi_E(t) = \sum a_i \tau^i \tag{5-1}$$

with $a_i \in R$ $a_i \to 0$, and $a_0 = \pi = \theta(t)$. For brevity, we will typically write the pair (E, φ_E) as E. We remind the reader that $\hat{\mathbb{G}}_a$ implicitly has the tautological A-module structure defined in Section 4C.

The main purpose of this section is to establish some facts that will be used in the proof of Theorem 6.2 below. We emphasize that in this application E will not be a Drinfeld module.

Proposition 5.1. *Any A-linear morphism $f : E \to G$ between admissible A-modules is determined by the induced morphism on tangent spaces. More precisely, if we write $\varphi_E(t) = \pi\tau^0 + \sum_{j\geq 1} a_j \tau^j$, $\varphi_G(t) = \pi\tau^0 + \sum_{j\geq 1} c_j \tau^j$, and $f = \sum_i b_i \tau^i$, then f is determined by b_0, as follows*:

$$b_r = \frac{1}{\pi - \pi^{q^r}} \sum_{i=0}^{r-1} (b_i a_{r-i}^{q^i} - c_{r-i} b_i^{q^{r-i}}).$$

Proof. Because f is B-linear, we have

$$\left(\sum_{i\geq 0} b_i\tau^i\right)\left(\pi\tau^0+\sum_{j\geq 1} a_j\tau^j\right)=\left(\pi\tau^0+\sum_{j\geq 1} c_j\tau^j\right)\left(\sum_{i\geq 0} b_i\tau^i\right).$$

Comparing the coefficients of τ^r, we have

$$b_0 a_r^{q^0}+\cdots+b_{r-1}a_1^{q^{r-1}}+b_r\pi^{q^r}=\pi b_r+c_1 b_{r-1}^{q}+\cdots+c_r b_0^{q^r}.$$

Therefore we have

$$b_r(\pi-\pi^{q^r})=\sum_{i=0}^{r-1}(b_i a_{r-i}^{q^i}-c_{r-i}b_i^{q^{r-i}}).$$

Since R is π-torsion free and $1-\pi^{q^r-1}$ is invertible for $r\geq 1$, this determines each b_r uniquely in terms of $b_0,\ldots,b_{r-1}$. Therefore b_0 determines each b_r. □

Corollary 5.2. *The R-module map $R\to \mathrm{Hom}_A(\hat{\mathbb{G}}_{\mathrm{a}},\hat{\mathbb{G}}_{\mathrm{a}})$ defined by $b\mapsto b\tau^0$ is an isomorphism.*

Now consider the subset $S^\dagger\subset R\{\tau\}\hat{}$ defined by

$$S^\dagger:=\left\{\sum_{i\geq 0} b_i\tau^i\in R\{\tau\}\hat{}\mid v(b_i)\geq i,\ \text{for all } i \text{ and } b_0\in R^*\right\}. \tag{5-2}$$

Here, and below, we write $v(b)$ for the minimal i such that $b\in\mathfrak{p}^i R$. (Note that v may not be a valuation if R is not a discrete valuation ring.)

Proposition 5.3. *$S^\dagger$ is a group under composition.*

Note that a similar group of automorphisms appears in [Dupuy 2014, §4.3].

Proof. The fact that $S^\dagger$ is a submonoid of $R\{\tau\}\hat{}$ under composition follows immediately from the law $b\tau^i\circ c\tau^j=bc^{q^i}\tau^{i+j}$ and linearity. Indeed if $v(b)\geq i$ and $v(c)\geq j$, then $v(bc^{q^i})\geq i+j$.

Now let us show that any element $f=\sum b_i\tau^i\in S^\dagger$ has an inverse under composition. Let $g=\sum_{n=0}^\infty c_n\tau^n$, where $c_0=b_0^{-1}$ and we define inductively $c_n=-b_0^{-q^n}(c_0b_n+c_1b_{n-1}^q+\cdots+c_{n-1}b_1^{q^{n-1}})$. Then it is easy to check that $g\circ f=\mathbb{1}$. Take $n\geq 1$ and assume $v(c_i)\geq i$ for all $i=0,\ldots,n-1$. Then it is enough to show $v(c_n)\geq n$. We have $v(c_n)\geq\min\{v(c_ib_{n-i}^{q^i})\mid i=0,\ldots,n-1\}$. Now

$$v(c_ib_{n-i}^{q^i})=v(c_i)+q^iv(b_{n-i})=i+q^i(n-i)\geq i+(n-i)=n.$$

Therefore the left inverse g of f lies in $S^\dagger$.

Now consider $g'=\sum_{n=0}^\infty d_n\tau^n\in R\{\{\tau\}\}$, where $d_0=b_0^{-1}$ and we inductively define

$$d_n=-b_0^{-1}(b_1d_{n-1}^{q^1}+b_2d_{n-2}^{q^2}+\cdots+b_nd_0^{q^n}).$$

Then as above, one can easily check that $f\circ g'=\mathbb{1}$ and hence it is a right inverse of f in $R\{\{\tau\}\}$. But using the associativity property of $R\{\{\tau\}\}$ we get $g'=(g\circ f)\circ g'=g\circ(f\circ g')=g$ and hence g is both a left and right inverse of f in $S^\dagger$. □

Proposition 5.4. *Let B denote the subring $\mathbb{F}_q[t] \subseteq A$. Let $f : E \to G$ be a B-linear homomorphism of admissible A-module schemes over* Spf R. *Then f is A-linear.*

Proof. Given any element $a \in A$, we will show $\varphi_G(a) \circ f = f \circ \varphi_E(a)$. Both sides are B-linear homomorphisms $E \to G$; indeed, f is B-linear by assumption, and both $\varphi_G(a)$ and $\varphi_E(a)$ are B-linear because A is commutative. Furthermore, on tangent spaces, $\varphi_G(a) \circ f$ is multiplication by $af'(0)$, and $f \circ \varphi_E(a)$ is multiplication by $f'(0)a$; this is because the A-module schemes are admissible. Thus the two morphisms agree on tangent spaces and therefore they agree, by Proposition 5.1. □

In other words, the forgetful functor from admissible A-modules schemes over R to admissible B-module schemes over R is fully faithful. This remains true if we allow B to be not just $\mathbb{F}_q[t]$ but any sub-$\mathbb{F}_q$-algebra of A strictly containing $\mathbb{F}_q$.

Lemma 5.5. *If $q \geq 3$, then $q^i - q^{i-j} - j - 1 \geq 0$ for all $j = 1, \ldots, i$.*

Proof. Consider $f(x) = q^i - q^{i-x} - x - 1$, for $1 \leq x \leq i$. Then $f(1) \geq 0$ since $q \geq 3$. Now $f'(x) = q^{i-x} \ln q - 1$. Since $\ln q > 1$ for $q \geq 3$, we have $f'(x) \geq 0$ for all $1 \leq x \leq i$ and hence $f(x) \geq 0$ for all $1 \leq x \leq i$ and we are done. □

Lemma 5.6. *For $q = 2$ and $i \geq 2$, $q^i - i - 1 \geq 1$.*

Proof. Consider the function $h(x) = q^x - x$ for $x \geq 2$. Then $h'(x) = q^x \ln q - 1 = \ln q^{q^x} - 1 > 0$ since $x \geq 2$. Therefore h is a strictly increasing function and hence the minimum is attained at i=2. Therefore $q^i - i \geq q^2 - 2 = 2$ and the result follows. □

Lemma 5.7. *For $q = 2$ and $i \geq 2$ and $j = 1, \ldots, i$*

$$q^i - q^{i-j} - j \geq 1$$

Proof. For $j = i$, the result follows from Lemma 5.6. So we may assume $1 \leq j \leq i - 1$. Let $H(x) := q^i - q^{i-x} - x$ where $1 \leq x \leq i - 1$. Then $H'(x) = q^{i-x} \ln q - 1$. Since $x \leq i - 1$ implies $i - x \geq 1$ and hence $q^{i-x} \geq q$. Therefore we get

$$H'(x) \geq q \ln q - 1 = \ln q^q - 1 = \ln 4 - 1 > 0,$$

where the last equality follows since $q = 2$.

Hence $H(x)$ is a strictly increasing function within the interval $1 \leq x \leq i - 1$. Therefore the minimum is achieved at $x = 1$ and we have

$$\begin{aligned} q^i - q^{i-j} - j &\geq q^i - q^{i-1} - 1 \\ &= q^{i-1}(q-1) - 1 \\ &= q^{i-1} - 1 \qquad \text{(because } q = 2\text{)} \\ &\geq q - 1 \qquad \text{(since } i \geq 2\text{)} \\ &= 1 \end{aligned}$$

□

Theorem 5.8. *Suppose $v(a_i) \geq q^i - 1$ for all $i \geq 1$, where the a_i are as in (5-1). Then there exists a unique homomorphism $f : E \to \hat{\mathbb{G}}_a$ of A-module schemes over S, written $f = \sum_{i=0}^{\infty} b_i \tau^i$ in coordinates, with $b_0 = 1$. Moreover,*

(1) *if $q \geq 3$, then $v(b_i) \geq i$ and f is an isomorphism of A-module schemes;*

(2) *if $q = 2$, then $v(b_i) \geq i - 1$.*

Proof. Let $f = \sum_{i=0}^{\infty} b_i \tau^i$, $b_i \in R$, where $b_0 = 1$ and

$$b_i = \pi^{-1}(1 - \pi^{q^i - 1})^{-1} \sum_{j=1}^{i} b_{i-j} a_j^{q^{i-j}}. \tag{5-3}$$

Indeed, this is the only possible choice for f, by Corollary 5.2. Conversely, it is easy to see that f satisfies $\varphi(t) \circ f = f \circ \varphi(t)$, which implies $\varphi(b) \circ f = f \circ \varphi(b)$ for all $b \in B$.

(1) Assume $q \geq 3$. Let us now show $v(b_i) \geq i$. For $i = 0$, it is clear. For $i \geq 1$, we may assume by induction that $v(b_j) \geq j$ for all $j = 1, \ldots, i-1$. By (5-3), we have $v(b_i) \geq \min\{v(b_{i-j} a_j^{q^{i-j}}) - 1 \mid j = 1, \ldots, i\}$. Now

$$\begin{aligned} v(b_{i-j} a_j^{q^{i-j}}) - 1 &\geq v(b_{i-j}) + v(a_j^{q^{i-j}}) - 1 \\ &\geq i - j + q^{i-j}(q^j - 1) - 1 \\ &= i - j + q^i - q^{i-j} - 1 \\ &\geq i \qquad \text{(by Lemma 5.5).} \end{aligned}$$

Therefore we have $v(b_i) \geq i$.

Therefore f is a restricted power series and hence defines a map between π-formal schemes $f : E \to \hat{\mathbb{G}}_a$ which is A-linear.

Let us show that f is an isomorphism. By Proposition 5.3, there exists a linear map $g : \hat{\mathbb{G}}_a \to E$ such that $f \circ g = g \circ f = \mathbb{1}$. Then g is also A-linear for formal reasons: for any $a \in A$, we have $f(g(\varphi(a)x)) = \varphi(a)x = f(\varphi(a)g(x))$. Since f is injective, we must have $g(\varphi(a)x) = \varphi(a)g(x)$ which shows the A-linearity of g and we are done.

(2) Now assume $q = 2$. We want to show that $v(b_i) \geq i - 1$ for all $i \geq 1$. For $i = 1$, we have $b_1 = \pi^{-1}(1 - \pi^{q-1})^{-1}(b_0 a_1)$ and hence $v(b_1) \geq q - 1 - 1 = 0$. For $i \geq 2$ and $j = 1, \ldots, i$,

$$v(b_{i-j} a_j^{q^{i-j}}) = v(b_{i-j}) + q^{i-j} v(a_j) \geq (i - j - 1) + (q^i - q^{i-j}) \quad (\text{since } v(a_j) \geq q^j - 1).$$

Hence to show $v(b_i) \geq i - 1$, it is enough to show that $q^i - q^{i-j} - j \geq 0$ and that follows from Lemma 5.7. □

The remainder of this section consists of an interesting observation which will not however be used in this paper. Letting $\mathbb{G}_a^{\text{for}}$ denote the formal completion of $\hat{\mathbb{G}}_a$ along the identity section $\operatorname{Spf} R \to \hat{\mathbb{G}}_a$. Thus we have $\mathbb{G}_a^{\text{for}} = \operatorname{Spf} R[\![x]\!]$, where $R[\![x]\!]$ has the (π, x)-adic topology. We want to extend the A-action on

$\mathbb{G}_{\mathrm{a}}^{\mathrm{for}}$ to an action of $\hat{A}$:

$$\hat{A} \to \mathrm{End}_{\mathbb{F}_q}(\mathbb{G}_{\mathrm{a}}^{\mathrm{for}}/S). \tag{5-4}$$

Recall that $\mathrm{End}_{\mathbb{F}_q}(\mathbb{G}_{\mathrm{a}}^{\mathrm{for}})$ agrees with the noncommutative power-series ring $R\{\{\tau\}\}$, with commutation law $\tau b = b^q \tau$ for $b \in R$. (See for example [Drinfeld 1974, §2].) Therefore for any $a \in A$, we can write

$$\varphi(a) = \sum_j \alpha_j \tau^j$$

where $\alpha_j \in R$. Each α_j can be thought of as a function of $a \in A$. To construct (5-4) it is enough to prove that these functions are $\mathfrak{p}$-adically continuous, which also implies that such an extension to a continuous $\hat{A}$-action is unique. This is a consequence of the following result.

Proposition 5.9. *If $a \in \mathfrak{p}^n$, then $\alpha_j \in \mathfrak{p}^{n-j} R$.*

Proof. Clearly, it is true for $n = 0$. Now assume it is true for some given n. Suppose $a \in \mathfrak{p}^{n+1}$ and write $a = \pi b$, where $b \in \mathfrak{p}^n$. Let $\varphi(b) = \sum_j \beta_j \tau^j$ and $\varphi(\pi) = \sum_k \gamma_k \tau^k$. Then we have

$$\sum_j \alpha_j \tau^j = \varphi(a) = \varphi(\pi)\varphi(b) = \sum_k \gamma_k \tau^k \sum_j \beta_j \tau^j = \sum_{k,j} \gamma_k \beta_j^{q^k} \tau^{j+k}$$

and hence $\alpha_j = \sum_{k=0}^{j} \gamma_k \beta_{j-k}^{q^k}$. So to show $\alpha_j \in \mathfrak{p}^{n+1-j} R$, it suffices to show

$$\gamma_k \beta_{j-k}^{q^k} \in \mathfrak{p}^{n+1-j} R, \quad \text{for } 0 \le k \le j \le n+1.$$

By induction we have $\beta_{j-k} \in \mathfrak{p}^{n-(j-k)} R$ and hence $\gamma_k \beta_{j-k}^{q^k} \in \mathfrak{p}^{(n-(j-k))q^k} R$. Since we have $(n-(j-k))q^k \ge n - j + 1$ for $k \ge 1$, we then have $\gamma_k \beta_{j-k}^{q^k} \in \mathfrak{p}^{n-j+1} R$. For $k = 0$, because φ is a strict module structure, we have $\gamma_0 = \pi$ and hence $\gamma_0 \beta_j \in \pi \mathfrak{p}^{n-j} R = \mathfrak{p}^{1+n-j} R$. □

6. Characters of N^n — upper bounds

We continue to let E denote the admissible A-module scheme over S of (5-1). Let N^n denote the kernel of the projection $u : J^n E \to E$. Thus we have a short exact sequence of A-module schemes over S:

$$0 \to N^n \to J^n E \xrightarrow{u} E \to 0$$

The purpose of this section is to analyze the character group of N^n. In the applications of this section, E will eventually be a Drinfeld module, but we do not need to assume this yet.

Let us fix a coordinate x on E, and denote the corresponding Buium–Joyal coordinates on $J^n E$ by $x, x', \ldots, x^{(n)}$. From now on, let us abusively write ϕ for the Frobenius pull back ϕ^* of (4-6).

Lemma 6.1. *For all $n \ge 0$, $\phi^n(x) = \pi^n x^{(n)} + O(n-1)$, where $O(n-1)$ are elements of order less than or equal to $n-1$.*

Proof. For $n = 0$, it is clear. For $n \geq 1$, we have by induction

$$\begin{aligned}\phi^n(x) &= \phi(\pi^{n-1}x^{(n-1)} + O(n-2)) \\ &= \pi^{n-1}\phi(x^{(n-1)}) + O(n-1) \\ &= \pi^{n-1}(\pi\delta(x^{(n-1)}) + (x^{(n-1)})^{\hat{q}}) + O(n-1) \\ &= \pi^n x^{(n)} + O(n-1).\end{aligned}$$

□

Theorem 6.2. *For any $n \geq 1$, let H^n denote the kernel of the projection $u : J^n E \to J^{n-1}E$. Then there is a unique A-linear homomorphism $\vartheta_n : H^n \to \hat{\mathbb{G}}_a$ of the form*

$$\vartheta_n(x^{(n)}) = x^{(n)} + b_1(x^{(n)})^q + b_2(x^{(n)})^{q^2} + \cdots,$$

where $b_i \in R$. Moreover, ϑ_n freely generates $\mathrm{Hom}_A(H^n, \hat{\mathbb{G}}_a)$ as an R-module, and

(1) *if $q \geq 3$, then $v(b_i) \geq i$ and ϑ_n is an isomorphism of A-module schemes;*

(2) *if $q = 2$, then $v(b_i) \geq i - 1$.*

Proof. First observe that we have

$$\varphi_E(t)\phi^n(x) = \phi^n(\varphi_E(t)) = \phi^n(\pi)\phi^n(x) + \phi^n(a_1)\phi^n(x)^q + \cdots + \phi^n(a_r)\phi^n(x)^{q^r}.$$

Second, the subscheme H^n is defined by setting the $x, x', \ldots, x^{(n-1)}$ coordinates to 0. Combining these two observations and Lemma 6.1, we obtain

$$\pi^n\varphi_E(t)x^{(n)} = \pi\pi^n x^{(n)} + \phi^n(a_1)(\pi^n x^{(n)})^q + \cdots + \phi^n(a_r)(\pi^n x^{(n)})^{q^r}$$

and hence

$$\varphi_E(t)x^{(n)} = \pi x^{(n)} + \phi^n(a_1)\pi^{n(q-1)}(x^{(n)})^q + \cdots + \phi^n(a_r)\pi^{n(q^r-1)}(x^{(n)})^{q^r}.$$

But then by Theorem 5.8, there is a unique A-linear homomorphism ϑ_n of the kind desired for the respective cases of $q \geq 3$ and $q = 2$. Moreover by Proposition 5.1, $\mathrm{Hom}_A(H^n, \hat{\mathbb{G}}_a)$ is freely generated by ϑ_n as an R-module. Finally, by Proposition 5.3, ϑ_n an isomorphism when $q \geq 3$. □

Now consider the exact sequence

$$0 \to H^n \to N^n \to N^{n-1} \to 0$$

and the corresponding long exact sequence

$$0 \to \mathrm{Hom}_A(N^{n-1}, \hat{\mathbb{G}}_a) \to \mathrm{Hom}_A(N^n, \hat{\mathbb{G}}_a) \to \mathrm{Hom}_A(H^n, \hat{\mathbb{G}}_a) \to \cdots.$$

The image of the map $\mathrm{Hom}_A(N^n, \hat{\mathbb{G}}_a) \to \mathrm{Hom}_A(H^n, \hat{\mathbb{G}}_a)$ can be regarded as a sub-R-module of R, by Theorem 6.2 above. Therefore in the R-module filtration

$$\mathrm{Hom}_A(N^n, \hat{\mathbb{G}}_a) \supseteq \mathrm{Hom}_A(N^{n-1}, \hat{\mathbb{G}}_a) \supseteq \cdots \supseteq \mathrm{Hom}_A(N^0, \hat{\mathbb{G}}_a) = 0,$$

each associated graded piece is canonically a submodule of R.

In particular, we have the following:

Proposition 6.3. *If R is a discrete valuation ring, then $\mathrm{Hom}_A(N^n, \hat{\mathbb{G}}_{\mathrm{a}})$ is a free R-module of rank at most n.*

7. The lateral Frobenius and characters of N^n

We continue to let E denote the admissible A-module scheme over S of (5-1).

Now we will construct a family of important operators which we call the *lateral Frobenius* operators. That is, for all n, we will construct maps $\mathfrak{f} : N^{n+1} \to N^n$ which are lifts of Frobenius relative to the projections $u : N^{n+1} \to N^n$ and hence make the system $\{N^n\}_{n=0}^{\infty}$ into a prolongation sequence. Do note that *a priori* the A-modules N^n do not form a prolongation sequence to start with.

Let N^∞ denote the inverse limit the projection maps $u : N^{n+1} \to N^n$. (Here and below, we take inverse limits in the category of presheaves on R-algebras in which π is nilpotent. They are representable by affine formal schemes.) Then the maps $\mathfrak{f}$ induce a lift of Frobenius on N^∞. Similarly on $J^\infty E = \lim_n J^n E$, the maps ϕ induce a lift of Frobenius. Now for all $n \geq 1$, the inclusion $N^n \hookrightarrow J^n E$ is a closed immersion and hence induces a closed immersion of schemes $N^\infty \hookrightarrow J^\infty E$. But $\mathfrak{f}$ is not obtained by restricting ϕ to N^∞. In fact, ϕ does not even preserve N^∞. So $\mathfrak{f}$ is an interesting operator which is distinct from ϕ, although it does satisfy a certain relation with ϕ which we will explain below.

Here we would also like to remark that the lateral Frobenius can also be constructed in the mixed-characteristic setting of p-jet spaces of arbitrary schemes [Borger and Saha 2017b], but it is much more involved.

Let $F : W_n \to W_{n-1}$ and $V : W_{n-1} \to W_n$ denote the Frobenius and Verschiebung maps of Section 3C. Let us arrange them in the following diagram, although it does not commute.

$$\begin{array}{ccc} W_n & \xrightarrow{V} & W_{n+1} \\ {\scriptstyle F}\downarrow & & \downarrow{\scriptstyle F} \\ W_{n-1} & \xrightarrow{V} & W_n \\ & & \downarrow{\scriptstyle F} \\ & & W_{n-1} \end{array} \tag{7-1}$$

Rather the following is true

$$FFV = FVF. \tag{7-2}$$

Indeed, the operator FV is multiplication by $\pi = \theta(t)$, and F is a morphism of A-algebras.

We can reexpress this in terms of jet spaces using the natural identifications $J^n E \simeq W_n$ and $N^n \simeq W_{n-1}$. For jet spaces, let us switch to the notation $i := V$ and $\phi := F$ for the right column of (7-1). Then we define the *lateral Frobenius*

$$\mathfrak{f} : N^{n+1} \to N^n$$

simply to be the map $F : W_n \to W_{n-1}$ in left column. Thus (7-1) becomes the following:

$$\begin{array}{ccc} N^{n+1} & \xrightarrow{i} & J^{n+1}E \\ \downarrow{\scriptstyle \mathfrak{f}} & & \downarrow{\scriptstyle \phi} \\ N^{n} & \xrightarrow{i} & J^{n}E \\ & & \downarrow{\scriptstyle \phi} \\ & & J^{n-1}E \end{array} \tag{7-3}$$

Note again that this diagram is not commutative. However rewriting (7-2) in the above notation, we do have

$$\phi^{\circ 2} \circ i = \phi \circ i \circ \mathfrak{f}. \tag{7-4}$$

We emphasize that when we use the notation N^n, the A-module structure will always be understood to be the one that makes i an A-linear morphism. It should not be confused with the A-module structure coming by transport of structure from the isomorphism $N^n \simeq W_{n-1} = J^{n-1}E$ of group schemes.

We also emphasize that while i is a morphism of S-schemes, the vertical arrows ϕ and $\mathfrak{f}$ in the diagram above lie over the Frobenius endomorphism ϕ of S, rather than the identity morphism.

Lemma 7.1. *For any torsion-free R-algebra B, the map $FV : W_n(B) \to W_n(B)$ is injective.*

Proof. Since B is torsion free, the ghost map $W_n(B) \to B \times \cdots \times B$ is injective, and hence $W_n(B)$ is torsion free. The result then follows because FV is multiplication by π. □

Proposition 7.2. *The morphism $\mathfrak{f} : N^n \to N^{n-1}$ is A-linear.*

Proof. We want to show that for any $a \in A$, the two morphisms $N^{n+1} \to N^n$ given by $x \mapsto a\mathfrak{f}(x)$ and by $x \mapsto \mathfrak{f}(ax)$ are equal. Since the N^i are flat over R, it is enough to consider B-points x, where B is a π-torsion free R-algebra.

Since both ϕ and i are A-linear morphisms, so are ϕi and $\phi^2 i$. Therefore we have

$$\phi i(\mathfrak{f}(ax)) = \phi^2 i(ax) = a\phi^2 i(x) = a\phi i(\mathfrak{f}(x)) = \phi i(a\mathfrak{f}(x)).$$

Thus the points $\mathfrak{f}(ax)$ and $a\mathfrak{f}(x)$ of $N^n(B)$ become equal after the application of ϕi. Now translating from the notation of diagram (7-3) to that of diagram (7-1), we have two elements of $W_{n-1}(B)$ which become equal after applying FV. But since $FV = \pi$ and B is torsion free, Lemma 7.1 implies these two elements must be equal. □

For $0 \le i \le k-1$, let us abusively write $\mathfrak{f}^{\circ i}$ for the composition

$$\mathfrak{f}^{\circ i} : N^n \xrightarrow{\overbrace{\mathfrak{f}\circ\cdots\circ\mathfrak{f}}^{i \text{ times}}} N^{n-i} \xrightarrow{u} N^{n-k}.$$

Then for all $1 \le i \le n$, we define the *canonical characters* $\Psi_i \in \mathrm{Hom}_A(N^n, \hat{\mathbb{G}}_a)$ (associated to our implicit coordinate x on E) by

$$\Psi_i = \vartheta_1 \circ \mathfrak{f}^{\circ i-1} \tag{7-5}$$

where ϑ_1 is as in Theorem 6.2. Clearly, the maps Ψ_i are A-linear since each one of the maps above is. Finally, given a character $\Psi \in \mathrm{Hom}_A(N^{n-1}, \hat{\mathbb{G}}_{\mathrm{a}})$, we will write $\mathfrak{f}^*\Psi = \Psi \circ \mathfrak{f}$. Note that $\mathfrak{f}^*$ is semilinear: for $\lambda \in R$, we have

$$\mathfrak{f}^*(\lambda\Psi) = \phi(\lambda)\,\mathfrak{f}^*(\Psi). \tag{7-6}$$

The points of J^nE contained in N^n are those with Witt coordinates of the form $(0, x_1, x_2, \dots, x_n)$. We will use the abbreviated coordinates $(x_1, \dots, x_n)$ on N^n instead.

Lemma 7.3. *For all $i = 1, \dots, n$, we have*

$$\Psi_i(x_1, \dots, x_n) \equiv \begin{cases} x_1^{\hat{q}^{i-1}} \bmod \pi & \text{if } q \geq 3, \\ x_1^{\hat{q}^{i-1}} + \phi(a_1)x_1^{q\hat{q}^{i-1}} \bmod \pi & \text{if } q = 2, \end{cases}$$

where a_1 is the first of the structure constants of the Drinfeld module E, as in (5-1).

Proof. Since $\mathfrak{f}$ is identified with the Frobenius map $F : W_n \to W_{n-1}$, it reduces modulo π to the $\hat{q}$-th power of the projection map. Therefore, we have

$$\Psi_i(x_1, \dots, x_n) = \vartheta_1 \circ \mathfrak{f}^{\circ(i-1)}(x_1, \dots, x_n) \equiv \vartheta_1(x_1^{\hat{q}^{i-1}}) \bmod \pi.$$

$\underline{q \geq 3}$ By part (1) of Theorem 6.2, the map ϑ_1 is congruent to the identity modulo π. Therefore Ψ_i is congruent to $x_1^{\hat{q}^{i-1}}$ modulo π.

$\underline{q = 2}$ By part (2) of Theorem 6.2, we have $\vartheta_1(x_1) \equiv x_1 + b_1 x_1^q \bmod \pi$, where by (5-3), we have

$$b_1 = \pi^{-1}(1 - \pi^{q-1})^{-1}\pi^{q-1}\phi(a_1) \equiv \phi(a_1) \bmod \pi.$$

Therefore we have $\vartheta_1(x_1) \equiv x_1 + \phi(a_1)x_1^q \bmod \pi$, and so Ψ_i is congruent to $x_1^{\hat{q}^{i-1}} + \phi(a_1)x_1^{q\hat{q}^{i-1}} \bmod \pi$. □

Proposition 7.4. *If R is a discrete valuation ring, then the elements $\Psi_1, \dots, \Psi_n$ form an R-basis for* $\mathrm{Hom}_A(N^n, \hat{\mathbb{G}}_{\mathrm{a}})$.

Proof. By Proposition 6.3, the R-module $\mathrm{Hom}_A(N^n, \hat{\mathbb{G}}_{\mathrm{a}})$ is free of rank at most n. So to show the elements $\Psi_1, \dots, \Psi_n$ form a basis, it is enough by Nakayama's lemma to show they are linearly independent modulo π.

We can view $\mathrm{Hom}_A(N^n, \hat{\mathbb{G}}_{\mathrm{a}})$ as the set of additive functions in $\mathcal{O}(N^n)$. Further since N^n is flat, $\mathcal{O}(N^n)$ is π-torsion free, and so any function $f \in \mathcal{O}(N^n)$ is additive if πf is. Therefore the map $R/\pi R \otimes_R \mathrm{Hom}_A(N^n, \hat{\mathbb{G}}_{\mathrm{a}}) \to R/\pi R \otimes_R \mathcal{O}(N^n)$ remains injective.

So to show they are linearly independent in $R/\pi R \otimes_R \mathrm{Hom}_A(N^n, \hat{\mathbb{G}}_{\mathrm{a}})$, it is enough to show that $R/\pi R \otimes_R \mathrm{Hom}_A(N^n, \hat{\mathbb{G}}_{\mathrm{a}})$ maps injectively to $R/\pi R \otimes_R \mathcal{O}(N^n)$. Now by Lemma 7.3, we have $\Psi_i \equiv x_1^{\hat{q}^{i-1}} \bmod \pi$ for $q \geq 3$ (and $\Psi_i \equiv x_1^{\hat{q}^{i-1}} + \phi(a_1)x_1^{q\hat{q}^{i-1}}$ for $q = 2$). So the Ψ_i map to linearly independent elements of $R/\pi R \otimes_R \mathcal{O}(N^n)$. □

8. $\boldsymbol{X}_\infty(E)$

We now assume further that R is a discrete valuation ring and E is a Drinfeld module over Spf R. Let r denote the rank of E. We continue to write $\varphi_E(t) = a_0\tau^0 + a_1\tau^1 + \cdots + a_r\tau^r$, where $a_0 = \pi$, $a_i \in R$ for all i, and $a_r \in R^*$.

In this section and the next, we will determine the structure of $\boldsymbol{X}_\infty(E)$. In the case of elliptic curves, it falls in two distinct cases as to when the elliptic curve admits a lift of Frobenius and when not. In particular, canonical lifts of ordinary elliptic curves all fall into one case. A similar story happens in our case when E is a Drinfeld module of rank 2, which one might consider the closest analogue of an elliptic curve. However, when the rank exceeds 2, the behavior of $\boldsymbol{X}_\infty(E)$ offers much more interesting cases which leads us to introduce the concept of the *splitting order* m of a Drinfeld module E. The splitting order is always less than or equal to the rank of E. When the rank equals 2, the splitting order is 1 if and only if E admits a lift of Frobenius.

We would like to point out here that our structure result for $\boldsymbol{X}_\infty(E)$ is an integral version of the equal-characteristic analogue of [Buium 1995]. He shows that $\boldsymbol{X}_\infty(E) \otimes_R K$ is generated by a single element as a $K\{\phi^*\}$-module where $K = R\left[\frac{1}{p}\right]$. But here we show that the module $\boldsymbol{X}_\infty(E)$ itself is generated by a single element as a $R\{\phi^*\}$-module. These methods also work in the setting of elliptic curves over p-adic rings, and hence this stronger result can be achieved in that case too. (See [Borger and Saha 2017a].)

The following theorem should be viewed as an analogue of the fact that an elliptic curve has no nonzero homomorphism of $\mathbb{Z}$-module schemes to $\mathbb{G}_{\mathrm{a}}$. In our case, we show that no Drinfeld module admits a nonzero homomorphism of A-module schemes to $\hat{\mathbb{G}}_{\mathrm{a}}$.

Theorem 8.1. *We have* $\mathbf{X}_0(E) = \{0\}$.

Proof. Any character $f = \sum_{i\geq 0} b_i\tau^i \in \boldsymbol{X}_0(E)$ satisfies the following chain of equalities, where θ is as in (2-1):

$$\begin{aligned}
\varphi_{\hat{\mathbb{G}}_{\mathrm{a}}}(t) \circ f &= f \circ \varphi_E(t) \\
\theta(t)\tau^0 \circ \sum_{i\geq 0} b_i\tau^i &= \sum_{i\geq 0} b_i\tau^i \circ \sum_{j\geq 0} a_j\tau^j \\
\sum_{i\geq 0} \theta(t) b_i\tau^i &= \sum_{i\geq 0}\left(\sum_{j=0}^{r} b_{i-j} a_j^{q^{i-j}}\right)\tau^i
\end{aligned}$$

Comparing the coefficients of τ^i for $i > r$, and using the equality $a_0 = \theta(t)$, we have

$$b_i(1-\theta(t)^{q^i-1})\theta(t) = a_r^{q^{i-r}} b_{i-r} + a_{r-1}^{q^{i-r+1}} b_{i-r+1} + \cdots + a_1^{q^{i-1}} b_{i-1}. \tag{8-1}$$

Suppose f is nonzero. There exists an N such that $b_{N-r} \neq 0$ and $v(b_{N-r}) < v(b_i)$ for all $i \geq N-r+1$. Then the valuation of the right-hand side of (8-1), for $i = N$, becomes $v(a_r^{q^{i-r}} b_{N-r}) = v(b_{N-r})$, since

$v(a_r) = 0$. But then, by taking the valuation of both sides of (8-1), we have

$$v(b_N) = v(b_{N-r}) - 1 < v(b_{N-r})$$

and $N \geq N - r + 1$, which is a contradiction. Therefore f must be 0. □

As a consequence the short exact sequence of A-module schemes over S

$$0 \to N^n \xrightarrow{i} J^n E \to E \to 0, \tag{8-2}$$

induces an exact sequence

$$0 \to \boldsymbol{X}_n(E) \xrightarrow{i^*} \mathrm{Hom}_A(N^n, \hat{\mathbb{G}}_{\mathrm{a}}) \xrightarrow{\partial} \mathrm{Ext}_A(E, \hat{\mathbb{G}}_{\mathrm{a}}), \tag{8-3}$$

where $\mathrm{Ext}_A(E, \hat{\mathbb{G}}_{\mathrm{a}})$ denotes the group of extension classes of A-module schemes over R, as defined in Gekeler [1990a, §5]. He further defines an exact sequence

$$0 \to \mathrm{Lie}(E)^* \to \mathrm{Ext}_A^\sharp(E, \hat{\mathbb{G}}_{\mathrm{a}}) \to \mathrm{Ext}_A(E, \hat{\mathbb{G}}_{\mathrm{a}}) \to 0 \tag{8-4}$$

of R-modules, where $\mathrm{Ext}_A^\sharp(E, \hat{\mathbb{G}}_{\mathrm{a}})$ denotes the group of classes of an extension together with a splitting of the corresponding extension of Lie algebras. Finally one defines

$$\boldsymbol{H}_{\mathrm{dR}}(E) = \mathrm{Ext}_A^\sharp(E, \hat{\mathbb{G}}_{\mathrm{a}}). \tag{8-5}$$

Theorem 8.2. *The exact sequence (8-4) is split. The rank of* $\mathrm{Ext}_A(E, \hat{\mathbb{G}}_{\mathrm{a}})$ *is* $r - 1$, *and the rank of* $\mathrm{Ext}_A^\sharp(E, \hat{\mathbb{G}}_{\mathrm{a}})$ *is* r.

Proof. See Diagram (5.2) and Corollary 3.7 in [Gekeler 1990a]. □

The following is the equal-characteristic analogue of a result of Buium [1995, Proposition 3.2].

Theorem 8.3. *Let* (E, φ_E) *be a Drinfeld module of rank* r.

(1) $\boldsymbol{X}_r(E)$ *is nonzero.*

(2) *We have*

$$\boldsymbol{X}_1(E) \simeq \begin{cases} R & \text{if } E \text{ has a lift of Frobenius,} \\ \{0\} & \text{otherwise.} \end{cases}$$

Proof. (1) Consider the exact sequence (8-3). By Proposition 7.4, the R-module $\mathrm{Hom}_A(N^n, \hat{\mathbb{G}}_{\mathrm{a}})$ is free of rank n. But also $\mathrm{Ext}_A(E, \hat{\mathbb{G}}_{\mathrm{a}})$ is free of rank $r - 1$, by Theorem 8.2 above. Therefore when $n = r$, the kernel $\boldsymbol{X}_n(E)$ is nonzero.

(2) Now consider $\boldsymbol{X}_1(E)$. It is contained in $\mathrm{Hom}_A(N^1, \hat{\mathbb{G}}_{\mathrm{a}})$, which is free of rank 1, and the quotient is contained in $\mathrm{Ext}_A(E, \hat{\mathbb{G}}_{\mathrm{a}})$, which is torsion free. Therefore $\boldsymbol{X}_1(E)$ is either $\{0\}$ or all of $\mathrm{Hom}_A(N^1, \hat{\mathbb{G}}_{\mathrm{a}}) \simeq R$.

Let $\mathbb{1}$ denote the identity map in $\mathrm{Hom}_A(\hat{\mathbb{G}}_{\mathrm{a}}, \hat{\mathbb{G}}_{\mathrm{a}})$. Then its image $\partial(\mathbb{1})$ in $\mathrm{Ext}_A(E, \hat{\mathbb{G}}_{\mathrm{a}})$ is the class of the extension (8-2). Therefore we have the equivalences $\boldsymbol{X}_1(E) \simeq R \iff i^*$ is an isomorphism $\iff \partial(\mathbb{1}) = 0 \iff$ (8-2) is split $\iff E$ has a lift of Frobenius. □

Define the *splitting order* of the Drinfeld module E to be the integer m such that $\boldsymbol{X}_m(E) \neq \{0\}$ and $\boldsymbol{X}_{m-1}(E) = \{0\}$. We also say that E *splits at order* m. By Theorems 8.1 and 8.3 above, we have $1 \leq m \leq r$ and additionally $m = 1$ if and only if E has a Frobenius lift.

Computation of the character of the Carlitz module. Let $A = \mathbb{F}_q[t]$ with $q \geq 3$. Let E be the Carlitz module over R satisfying

$$\varphi_E(t)(x) = \pi x + x^q.$$

Then the operator $\varphi_E(t)$ itself is a lift of Frobenius and hence, by the universal property of J^1E, defines the A-linear splitting of the exact sequence

$$0 \to N^1 \to J^1E \to E \to 0$$

that is, an A-linear morphism $\nu : J^1E \to N^1$ given in Buium–Joyal coordinates by $\nu(x, x') = x' - x$. Then our normalized character $\Theta_1 : J^1E \to \hat{\mathbb{G}}_a$ is given by $\Theta_1 = \vartheta_1 \circ \nu$.

Define $L_i = (\pi^q - \pi)\cdots(\pi^{q^i} - \pi)$. Then from Theorem 6.2, we have $\vartheta_1 : N^1 \to \hat{\mathbb{G}}_a$ given by

$$\vartheta_1(x') = \frac{1}{\pi}\sum_{i=0}^{\infty} \frac{(-1)^i}{L_i}(\pi x')^{q^i}. \tag{8-6}$$

Hence we have

$$\Theta_1(x, x') = \frac{1}{\pi}\sum_{i=0}^{\infty} \frac{(-1)^i}{L_i}(\pi(x' - x))^{q^i} = \frac{1}{\pi}\log_C(\pi(x' - x)), \tag{8-7}$$

where $\log_C$ denotes the Carlitz logarithm, as in [Goss 1996, p. 57]. One can check that this is the exact analogue of Buium's character $\frac{1}{p}\log\big(1 + p\frac{x'}{x^p}\big)$ for $\hat{\mathbb{G}}_m$ in the mixed-characteristic setting.

8A. ***Splitting of $J^n(E)$.*** The exact sequence (8-2) is split by the Teichmüller section $v : E \to J^nE$, as defined in Section 3. We emphasize that v is only a morphism of $\mathbb{F}_q$-module schemes and is not a morphism of A-module schemes. Nevertheless, it induces an isomorphism

$$J^n(E) \xrightarrow{\sim} E \times N^n$$

of $\mathbb{F}_q$-module schemes. Therefore for any character $\Theta \in \boldsymbol{X}_n(E)$, we can write $\Theta = g_\Theta \oplus \Psi_\Theta$ or

$$\Theta(x_0, \ldots, x_n) = g_\Theta(x_0) + \Psi_\Theta(x_1, \ldots, x_n), \tag{8-8}$$

where $\Psi_\Theta = i^*\Theta \in \mathrm{Hom}_A(N^n, \hat{\mathbb{G}}_a)$ and $g_\Theta = v^*\Theta$. We call g_Θ the *Teichmüller component* of Θ. Note that because v is only $\mathbb{F}_q$-linear, g_Θ is also only $\mathbb{F}_q$-linear. It still can, however, be expressed as an additive restricted power series. On the other hand, the restriction Ψ_Θ of Θ to N^n does remain A-linear.

Now consider the morphism

$$(\phi \circ i - i \circ \mathfrak{f}) : N^{n+1} \to J^nE, \tag{8-9}$$

in the notation of (7-3). It is an A-linear morphism by Proposition 7.2.

Proposition 8.4. *There exists a morphism h (necessarily unique and A-linear) making the diagram*

$$\begin{array}{ccc} N^{n+1} & \xrightarrow{\phi\circ i - i\circ\mathfrak{f}} & J^nE \\ {\scriptstyle u}\downarrow & \nearrow_{h} & \\ N^1 & & \end{array}$$

commute. In coordinates, it has the form

$$h(x_1) = (\pi x_1, c_1 x_1^{\hat{q}}, c_2 x_1^{\hat{q}^2}, \ldots),$$

for some $c_j \in R$.

Proof. By (7-1)–(7-3), the first statement is equivalent to showing that for any R-algebra B, there exists a map $h : B \to W_n(B)$ such that

$$\begin{array}{ccc} W_n(B) & \xrightarrow{FV-VF} & W_n(B) \\ \downarrow & \nearrow_{h} & \\ B & & \end{array}$$

commutes, where the vertical map is the projection onto the zeroth component. Now for any $y \in W_{n-1}(B)$, we have

$$(FV - VF)(Vy) = FVVy - VFVy = \pi Vy - V(\pi y) = 0.$$

So such a function h exists.

To conclude that $h(x)$ is of the given form, we use a homogeneity argument. Let $(z_0, z_1, \ldots)$ denote the ghost components of $(x_0, x_1, \ldots)$. If interpret each x_j as an indeterminate of degree $\hat{q}^j$, then each z_j is a homogenous polynomial in the $x_0, \ldots, x_j$ of degree $\hat{q}^j$ and with coefficients in A: $z_1 = x_0^{\hat{q}} + \pi x_1$, and so on. Solving for x_j in terms of $z_0, \ldots, z_j$, we see that x_j is a homogenous polynomial in the $z_0, \ldots, z_j$ with coefficients in $A\left[\frac{1}{\pi}\right]$.

Now let $(y_0, y_1, \ldots)$ denote $(FV - VF)(x_0, x_1, \ldots)$, where $y_j \in R[x_0, \ldots, x_j]$. Then the ghost components of $(y_0, y_1, \ldots)$ are $(\pi z_0, 0, 0, \ldots) = (\pi x_0, 0, 0, \ldots)$. It follows that $y_0 = \pi x_0$. Further, by the above, y_j is an element of $R[x_0, \ldots, x_j]$ but also a homogeneous polynomial in πx_0 of degree $\hat{q}^j$ and with coefficients in $A\left[\frac{1}{\pi}\right]$. Therefore it is of the form $c_j x_0^{\hat{q}^j}$ for some $c_j \in R$. □

Proposition 8.5. *Let* Θ *be a character in* $\boldsymbol{X}_n(E)$.

(1) *We have*

$$i^*\phi^*\Theta = \mathfrak{f}^*(i^*\Theta) + \gamma\Psi_1,$$

where $\gamma = \pi g'_\Theta(0)$ *and* $g'_\Theta(x_0)$ *denotes the usual derivative of the polynomial* $g_\Theta(x_0) \in R[x_0]$ *of* (8-8).

(2) *For* $n \geq 1$, *we have*

$$i^*(\phi^{\circ n})^*\Theta = (\mathfrak{f}^{n-1})^* i^*\phi^*\Theta.$$

Proof. (1) By Proposition 8.4, we have

$$(\phi \circ i - i \circ \mathfrak{f})(x_1, \ldots, x_{n+1}) = (\pi x_1, c_1 x_1^{\hat{q}}, c_2 x_1^{\hat{q}^2}, \ldots),$$

where $c_j \in R$. Therefore we have

$$((i^*\phi^* - \mathfrak{f}^* i^*)\Theta)(x_1, \ldots, x_{n+1}) = \Theta(\pi x_1, c_1 x_1^{\hat{q}}, \ldots) = g_\Theta(\pi x_1) + \Psi_\Theta(c_1 x_1^{\hat{q}}, \ldots). \tag{8-10}$$

In particular, the character $(i^*\phi^* - \mathfrak{f}^* i^*)\Theta$ depends only on x_1. Therefore it is of the form $\gamma\Psi_1$, for some $\gamma \in R$. Further since by Theorem 6.2 we have $\Psi_1'(0) = 1$, the coefficient γ is simply the linear coefficient of $(i^*\phi^* - \mathfrak{f}^* i^*)\Theta$, which by (8-10) is $\pi g'_\Theta(0)$.

(2) This is another way of expressing $\phi^{\circ n} \circ i = \phi \circ i \circ \mathfrak{f}^{\circ(n-1)}$, which follows from (7-4) by induction. □

8B. ***Frobenius and the filtration by order.*** We would like to fix a notational convention here. Let $u : J^n E \to J^{n'} E$ denote the canonical projection map for any $n' < n$, given in Witt coordinates by $u(x_0, \dots, x_n) = (x_0, \dots, x_{n'})$.

Consider the following morphism of exact sequences of A-modules

$$\begin{array}{ccccccccc} 0 & \longrightarrow & N^n & \xrightarrow{i} & J^n E & \xrightarrow{u} & E & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle u} & & \downarrow{\scriptstyle u} & & \| & & \\ 0 & \longrightarrow & N^{n-1} & \xrightarrow{i} & J^{n-1} E & \xrightarrow{u} & E & \longrightarrow & 0. \end{array}$$

Since $\boldsymbol{X}_0(E) = \{0\}$ by Theorem 8.1, applying $\mathrm{Hom}_A(-, \hat{\mathbb{G}}_{\mathrm{a}})$ to the above, we obtain the following morphism of exact sequences of R-modules

$$\begin{array}{ccccccc} 0 & \longrightarrow & \boldsymbol{X}_n(E) & \xrightarrow{i^*} & \mathrm{Hom}_A(N^n, \hat{\mathbb{G}}_{\mathrm{a}}) & \xrightarrow{\partial} & \mathrm{Ext}_A(E, \hat{\mathbb{G}}_{\mathrm{a}}) \\ & & \uparrow{\scriptstyle u^*} & & \uparrow{\scriptstyle u^*} & & \| \\ 0 & \longrightarrow & \boldsymbol{X}_{n-1}(E) & \xrightarrow{i^*} & \mathrm{Hom}_A(N^{n-1}, \hat{\mathbb{G}}_{\mathrm{a}}) & \xrightarrow{\partial} & \mathrm{Ext}_A(E, \hat{\mathbb{G}}_{\mathrm{a}}). \end{array}$$

Proposition 8.6. *For any $n \geq 0$, the diagram*

$$\begin{array}{ccc} \boldsymbol{X}_n(E)/\boldsymbol{X}_{n-1}(E) & \xrightarrow{\phi^*} & \boldsymbol{X}_{n+1}(E)/\boldsymbol{X}_n(E) \\ \downarrow{\scriptstyle i^*} & & \downarrow{\scriptstyle i^*} \\ \mathrm{Hom}_A(N^n, \hat{\mathbb{G}}_{\mathrm{a}})/\mathrm{Hom}_A(N^{n-1}, \hat{\mathbb{G}}_{\mathrm{a}}) & \xrightarrow[\sim]{\mathfrak{f}^*} & \mathrm{Hom}_A(N^{n+1}, \hat{\mathbb{G}}_{\mathrm{a}})/\mathrm{Hom}_A(N^n, \hat{\mathbb{G}}_{\mathrm{a}}) \end{array}$$

is commutative. The morphisms i^* and ϕ^* are injective, and $\mathfrak{f}^*$ is bijective.

In fact, we will show in Corollary 9.9 that all the morphisms in the diagram of Proposition 8.6 are isomorphisms.

Proof. For $n \geq 1$, commutativity of the diagram follows from Proposition 8.5; for $n = 0$, it follows from Theorem 8.1.

The maps i^* are injective because the projections $J^n E \to J^{n-1} E$ and $N^n \to N^{n-1}$ have the same kernel, and $\mathfrak{f}^*$ is an isomorphism by Proposition 7.4. It follows that ϕ^* is an injection. □

8C. ***The character $\boldsymbol{\Theta_m}$.*** Recall the exact sequence (8-3)

$$0 \to \boldsymbol{X}_n(E) \xrightarrow{i^*} \mathrm{Hom}_A(N^n, \hat{\mathbb{G}}_{\mathrm{a}}) \xrightarrow{\partial} \mathrm{Ext}_A(E, \hat{\mathbb{G}}_{\mathrm{a}}).$$

Let m denote the splitting order of E. Then for all $n < m$, the map

$$\partial : \mathrm{Hom}_A(N^n, \hat{\mathbb{G}}_{\mathrm{a}}) \to \mathrm{Ext}_A(E, \hat{\mathbb{G}}_{\mathrm{a}})$$

is injective since $\boldsymbol{X}_n(E) = \{0\}$. But at $n = m$, we have $\boldsymbol{X}_m(E) \neq \{0\}$, and so there is a nonzero character $\Psi \in \mathrm{Hom}_A(N^m, \hat{\mathbb{G}}_{\mathrm{a}})$ in the kernel of ∂. Write Ψ in terms of the basis of canonical characters Ψ_i defined in (7-5):

$$\Psi = \tilde{\lambda}_m \Psi_m - \tilde{\lambda}_{m-1}\Psi_{m-1} - \cdots - \tilde{\lambda}_1 \Psi_1,$$

where $\tilde{\lambda}_i \in R$ for all $i = 0, \ldots, m-1$. Then we necessarily have $\tilde{\lambda}_m \neq 0$ since $\boldsymbol{X}_{m-1} = \{0\}$. Therefore we have

$$\partial \Psi_m = \lambda_{m-1}\partial\Psi_{m-1} + \cdots + \lambda_1 \partial \Psi_1 \in \mathrm{Ext}_A(E, \hat{\mathbb{G}}_{\mathrm{a}})_K \tag{8-11}$$

where $\lambda_i = \tilde{\lambda}_i/\tilde{\lambda}_m$ for all $i = 1, \ldots, m-1$. This implies that the character

$$\Psi_m - \lambda_1 \Psi_{m-1} - \cdots - \lambda_{m-1}\Psi_1$$

is in $\ker(\partial)$ and hence by the main exact sequence (8-3), there exists a unique $\Theta_m \in \boldsymbol{X}_m(E)_K$ such that

$$i^*\Theta_m = \Psi_m - \lambda_{m-1}\Psi_{m-1} - \cdots - \lambda_1 \Psi_1. \tag{8-12}$$

It then follows immediately that Θ_m is a K-linear basis for $\boldsymbol{X}_m(E)_K$, say by Propositions 7.4 and 8.6. (We will show in Corollary 9.9 that Θ_m actually lies in the group $\boldsymbol{X}_m(E)$ of integral characters, and is in fact an integral basis for it.)

Proposition 8.7. *Let m denote the splitting order of E. Then for any $j \geq 0$, the character $i^*(\phi^*)^j\Theta_m$ agrees with Ψ_{m+j} modulo rational characters of lower order, and the elements $\Theta_m, \phi^*\Theta_m, \cdots, \phi^{n-m*}\Theta_m$ are a basis of the K-vector space $\boldsymbol{X}_n(E)_K$.*

Proof. By Proposition 8.6, each character ${\phi^i}^*\Theta_m$ lies in $\boldsymbol{X}_{m+i}(E)$ but not in $\boldsymbol{X}_{m+i-1}(E)$. Therefore they are linearly independent. In particular, the rank of $\boldsymbol{X}_n(E)$ is at least $n - m + 1$.

At the same time, by Proposition 8.6, each $\boldsymbol{X}_{m+i}(E)/\boldsymbol{X}_{m+i-1}(E)$ has rank at most 1. Thus the rank of $\boldsymbol{X}_n(E)$ actually equals $n - m + 1$, and so the elements in question form a K-basis of $\boldsymbol{X}_n(E)_K$. □

Do note that this result will be improved to an integral version in Theorem 9.10.

9. Ext groups and de Rham cohomology

We will prove Theorem 1.1 in this section. We continue with the notation from the previous section. In particular, R is a discrete valuation ring.

We will briefly describe our strategy in the next few lines. Recall from (8-12) the equality

$$i^*\Theta_m = \Psi_m - \lambda_{m-1}\Psi_{m-1} - \cdots - \lambda_1\Psi_1$$

where $\lambda_j \in K$. *A priori*, the elements λ_j need not belong to R, but we prove in Theorem 9.8 that they actually do. This implies that $i^*\Theta_m$ lies in $\mathrm{Hom}_A(N^m, \hat{\mathbb{G}}_{\mathrm{a}})$ and $\ker(\partial)$, and hence by the exact sequence

(8-3), we have $\Theta_m \in \boldsymbol{X}_m(E)$ — that is, the character Θ_m is integral. From there, it is an easy consequence that $\boldsymbol{X}_n(E)$ is generated by $\Theta_m, \ldots, \Theta_m^{\phi^{n-m}}$ as an R-module.

To prove Theorem 9.8, which says that all λ_j belong to R, requires some preparation. For all $n \geq 1$, we will define maps from $\mathrm{Hom}_A(N^n, \hat{\mathbb{G}}_{\mathrm{a}})$ to $\mathrm{Ext}^\sharp(E, \hat{\mathbb{G}}_{\mathrm{a}})$ which is also interpreted as the de Rham cohomology from associated to the Drinfeld module E. These maps are obtained by push-outs of $J^n E$ by $\Psi \in \mathrm{Hom}_A(N^n, \hat{\mathbb{G}}_{\mathrm{a}})$. To give an idea, do note that, for every $n \geq 1$, there are canonical elements $E^*_\Psi \in \mathrm{Ext}_A(E, \hat{\mathbb{G}}_{\mathrm{a}})$ where the E^*_Ψ is a push-out of $J^n E$ by Ψ as follows

$$\begin{array}{ccccccccc} 0 & \longrightarrow & N^n & \xrightarrow{i} & J^n E & \xrightarrow{u} & E & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle \Psi} & & \downarrow{\scriptstyle e_\Psi} & & \| & & \\ 0 & \longrightarrow & \hat{\mathbb{G}}_{\mathrm{a}} & \longrightarrow & E^*_\Psi & \longrightarrow & E & \longrightarrow & 0 \end{array}$$

as $E^*_\Psi \in \mathrm{Ext}_A(E, \hat{\mathbb{G}}_{\mathrm{a}})$. It leads to a very interesting theory of δ-modular forms over the moduli space of Drinfeld modules and will be studied in a subsequent paper. And similar to previous cases, the main principles carry over to the case of elliptic curves or abelian schemes as well.

Now we introduce the theory of extensions of A-module group schemes. Given an extension $\eta_C \in \mathrm{Ext}(G, T)$ and $f : T \to T'$ where G, T and T' are A-modules and f is an A-linear map we have the following diagram of the push-forward extension $f_* C$.

$$\begin{array}{ccccccccc} 0 & \longrightarrow & T & \longrightarrow & C & \longrightarrow & G & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle f} & & \downarrow & & \| & & \\ 0 & \longrightarrow & T' & \longrightarrow & f_* C & \longrightarrow & G & \longrightarrow & 0 \end{array}$$

The class of $f_* C$ is obtained as follows — the class of η_C is represented by a linear (not necessarily A-linear) function $\eta_C : G \to T$. Then $\eta_{f_* C}$ is represented by the class $\eta_{f_* C} = [f \circ \eta_C] \in \mathrm{Ext}(E, T')$. In terms of the action of $t \in A$, $\varphi_C(t)$ is given by $\left(\begin{smallmatrix} \varphi_G(t) & 0 \\ \eta_C & \varphi_T(t) \end{smallmatrix}\right)$ where $\eta_C : G \to T$. Then $\varphi_{f_* C}(t)$ is given by

$$\begin{pmatrix} \varphi_G(t) & 0 \\ f(\eta_C) & \varphi_{T'}(t) \end{pmatrix} \tag{9-1}$$

Now consider the exact sequence

$$0 \to N^n \xrightarrow{i} J^n E \xrightarrow{u} E \to 0 \tag{9-2}$$

Given a $\Psi \in \mathrm{Hom}_A(N^n, \hat{\mathbb{G}}_{\mathrm{a}})$ consider the push out

$$\begin{array}{ccccccccc} 0 & \longrightarrow & N^n & \xrightarrow{i} & J^n E & \xrightarrow{u} & E & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle \Psi} & & \downarrow{\scriptstyle e_\Psi} & & \| & & \\ 0 & \longrightarrow & \hat{\mathbb{G}}_{\mathrm{a}} & \xrightarrow{i} & E^*_\Psi & \longrightarrow & E & \longrightarrow & 0 \end{array}$$

where $E^*_\Psi = (J^nE \times \hat{\mathbb{G}}_a)/\,\Gamma(N^n)$ and $\Gamma(N^n) = \{(i(z), -\Psi(z)) \mid z \in N^n\} \subset J^nE \times \hat{\mathbb{G}}_a$ and $e_\Psi(x) = [x, 0] \in E^*_\Psi$.

The Teichmüller section $v : E \to J^n(E)$ is an $\mathbb{F}_q$-linear splitting of the sequence (9-2). The induced retraction

$$\rho = \mathbb{1} - v \circ u : J^n(E) \to N^n$$

is given in coordinates simply by $\rho : (x_0, \dots, x_n) \mapsto (x_1, \dots, x_n)$. Let us denote by s_{Witt} the morphism on Lie algebras induced by ρ. Thus we have the following split exact sequence of R-modules

$$0 \longrightarrow \operatorname{Lie} N^n \underset{s_{\mathrm{Witt}}}{\overset{Di}{\rightleftarrows}} \operatorname{Lie} J^nE \xrightarrow{Du} \operatorname{Lie}(E) \longrightarrow 0.$$

Let s_Ψ denote the induced splitting of the push out extension

$$0 \longrightarrow \operatorname{Lie} \hat{\mathbb{G}}_a \underset{s_\Psi}{\rightleftarrows} \operatorname{Lie}(E^*_\Psi) \longrightarrow \operatorname{Lie}(E) \longrightarrow 0.$$

It is given explicitly by $\tilde{s}_\Psi : \operatorname{Lie} J^nE \times \operatorname{Lie} \hat{\mathbb{G}}_a \to \operatorname{Lie} \hat{\mathbb{G}}_a$

$$\tilde{s}_\Psi(x, y) := D\Psi(s_{\mathrm{Witt}}(x)) + y$$

and

$$s_\Psi : \operatorname{Lie}(E^*_\Psi) = \frac{\operatorname{Lie} J^nE \times \operatorname{Lie} \hat{\mathbb{G}}_a}{\operatorname{Lie} \Gamma(N^n)} \to \operatorname{Lie} \hat{\mathbb{G}}_a.$$

Recall that $\operatorname{Ext}^\sharp(E, \hat{\mathbb{G}}_a)$ consists of an extension of A-module schemes together with a splitting s of the corresponding extension of Lie algebras. (See (8-4) above or [Gekeler 1990a, §5].) Therefore we have the following morphism of exact sequences

$$\begin{array}{ccccccc} 0 & \longrightarrow & \boldsymbol{X}_n(E) & \longrightarrow & \operatorname{Hom}_A(N^n, \hat{\mathbb{G}}_a) & \longrightarrow & \operatorname{Ext}(E, \hat{\mathbb{G}}_a) \\ & & \downarrow & & \downarrow{\scriptstyle \Psi \mapsto (E^*_\Psi, s_\Psi)} & & \| \\ 0 & \longrightarrow & \operatorname{Lie}(E)^* & \longrightarrow & \operatorname{Ext}^\sharp(E, \hat{\mathbb{G}}_a) & \longrightarrow & \operatorname{Ext}(E, \hat{\mathbb{G}}_a) \longrightarrow 0. \end{array} \tag{9-3}$$

Proposition 9.1. *Let Θ be a character in $\boldsymbol{X}_n(E)$, and put $\tilde{\Theta} = \phi^*\Theta$.*

(1) *The map $\boldsymbol{X}_n(E) \to \operatorname{Lie}(E)^*$ of (9-3) sends Θ to $-Dg_\Theta$.*

(2) *We have $g_{\tilde{\Theta}}(x) = g_\Theta(x^{\hat{q}})$ and*

$$\Psi_{\tilde{\Theta}}(y) = \Psi_\Theta(\rho(\phi(i(y)))) + g_\Theta(\pi y_1).$$

Proof. (1) Let us recall in explicit terms how the map is given. For the split extension $E \times \hat{\mathbb{G}}_a$, the retractions $\operatorname{Lie}(E) \times \operatorname{Lie} \hat{\mathbb{G}}_a = \operatorname{Lie}(E \times \hat{\mathbb{G}}_a) \to \operatorname{Lie} \hat{\mathbb{G}}_a$ are in bijection with maps $\operatorname{Lie}(E) \to \operatorname{Lie} \hat{\mathbb{G}}_a$, a retraction s corresponding to map $x \mapsto s(x, 0)$. Therefore to determine the image of Θ, we need to identify $E^*_{\Psi_\Theta}$ with a split extension and then apply this map to s_{Ψ_Θ}.

A trivialization of the extension $E^*_{\Psi_\Theta}$ is given by the map

$$\frac{J^n E \times \hat{\mathbb{G}}_a}{\Gamma(N^n)} = E^*_{\Psi_\Theta} \xrightarrow{\sim} E \times \hat{\mathbb{G}}_a$$

defined by $[a, b] \mapsto (u(a), \Theta(a) + b)$. The inverse isomorphism H is then given by the expression

$$H(x, y) = [v(x), y - \Theta(v(x))],$$

and so the composition $E \to E \times \hat{\mathbb{G}}_a \to E^*_{\Psi_\Theta} \to \hat{\mathbb{G}}_a$ is simply $-\Theta \circ v = -g_\Theta$, which induces the map $-Dg_\Theta$ on the Lie algebras.

(2) We have

$$\tilde{\Theta}(x) = \Theta(\phi(x)) = \Psi_\Theta(\rho(\phi(x))) + g_\Theta(x_0^{\hat{q}} + \pi x_1) = (\Psi_\Theta(\rho(\phi(x))) + g_\Theta(\pi x_1)) + g_\Theta(x_0^{\hat{q}}).$$

In other words, we have $\tilde{\Psi}(\rho(x)) = \Psi_\Theta(\rho(\phi(x))) + g_\Theta(\pi x_1)$ and $\tilde{g}(x_0) = g_\Theta(x_0^{\hat{q}})$. Setting $x = i(y)$, we obtain the desired result. □

Proposition 9.2. *If $\Psi \in i^*\phi^*(\boldsymbol{X}_n(E))$, then the class $(E^*_\Psi, s_\Psi) \in \mathrm{Ext}^\sharp(E, \hat{\mathbb{G}}_a)$ is zero.*

Proof. Write $\Psi = i^*\phi^*\Theta$. We know from diagram (9-3) that E^*_Ψ is a trivial extension since Ψ lies in $i^*\boldsymbol{X}_{n+1}(E)$. Now by part (2) of Proposition 9.1, we have $g_{\phi^*\Theta}(x_0) = g_\Theta(x_0^{\hat{q}})$ and hence $Dg_{\phi^*\Theta} = 0$. Therefore by part (1) of that proposition, the class in $\mathrm{Ext}^\sharp(E, \hat{\mathbb{G}}_a)$ is zero. □

9A. ***The F-crystal $\boldsymbol{H}(E)$.*** The ϕ-linear map $\phi^* : \boldsymbol{X}_{n-1}(E) \to \boldsymbol{X}_n(E)$ induces a linear map $\boldsymbol{X}_{n-1}(E)' \to \boldsymbol{X}_n(E)$, which we will abusively also denote ϕ^*. Here, for any R-module M, we write M' for its base change $R \otimes_{\phi, R} M$ via $\phi : R \to R$. We then define

$$\boldsymbol{H}_n(E) = \frac{\mathrm{Hom}_A(N^n, \hat{\mathbb{G}}_a)}{i^*\phi^*(\boldsymbol{X}_{n-1}(E)')}.$$

Then $u : N^{n+1} \to N^n$ induces $u^* : \mathrm{Hom}_A(N^n, \hat{\mathbb{G}}_a) \to \mathrm{Hom}_A(N^{n+1}, \hat{\mathbb{G}}_a)$. And since $u^*i^*\phi^*(\boldsymbol{X}_n(E)) = i^*u^*\phi^*(\boldsymbol{X}_n(E)) = i^*\phi^*u^*(\boldsymbol{X}_n(E)) \subset i^*\phi^*(\boldsymbol{X}_{n+1}(E))$, it also induces a map $u^* : \boldsymbol{H}_n(E) \to \boldsymbol{H}_{n+1}(E)$. Define

$$\boldsymbol{H}(E) = \varinjlim \boldsymbol{H}_n(E), \tag{9-4}$$

where the limit is taken in the category of R-modules.

Similarly, $\mathfrak{f} : N^{n+1} \to N^n$ induces $\mathfrak{f}^* : \mathrm{Hom}_A(N^n, \hat{\mathbb{G}}_a) \to \mathrm{Hom}_A(N^{n+1}, \hat{\mathbb{G}}_a)$, which descends to a ϕ-linear morphism of R-modules

$$\mathfrak{f}^* : \boldsymbol{H}_n(E) \to \boldsymbol{H}_{n+1}(E)$$

because we have $\mathfrak{f}^*i^*\phi^*(\boldsymbol{X}_{n-1}(E)) = i^*\phi^*\phi^*(\boldsymbol{X}_{n-1}(E) \subset i^*\phi^*\boldsymbol{X}_n(E)$. This then induces a ϕ-linear endomorphism $\mathfrak{f}^* : \boldsymbol{H}(E) \to \boldsymbol{H}(E)$.

Finally, let $\boldsymbol{I}_n(E)$ denote the image of ∂:

$$\boldsymbol{I}_n(E) = \mathrm{im}[\mathrm{Hom}(N^n, \hat{\mathbb{G}}_a) \xrightarrow{\partial} \mathrm{Ext}_A(E, \hat{\mathbb{G}}_a)]. \tag{9-5}$$

So $\mathrm{Hom}(N^n, \hat{\mathbb{G}}_{\mathrm{a}})/\boldsymbol{X}_n(E) \simeq \boldsymbol{I}_n(E)$. Then u induces maps $u^* : \boldsymbol{I}_n(E) \to \boldsymbol{I}_{n+1}(E)$, and we put

$$\boldsymbol{I}(E) = \varinjlim \boldsymbol{I}_n(E), \tag{9-6}$$

where again the limit is taken in the category of R-modules.

Proposition 9.3. *The morphism*

$$u^* : \boldsymbol{H}_n(E) \otimes K \to \boldsymbol{H}_{n+1}(E) \otimes K$$

is injective. For $n \geq m$, it is an isomorphism.

Proof. Consider the following diagram of exact sequences:

$$\begin{array}{ccccccccc}
 & & 0 & & 0 & & & & \\
 & & \uparrow & & \uparrow & & & & \\
 & & K\langle \phi^{\circ(n-m)*}\Theta\rangle' & \xrightarrow{i^*\phi^*} & K\langle \Psi_{n+1}\rangle & & & & \\
 & & \uparrow & & \uparrow & & & & \\
0 & \longrightarrow & \boldsymbol{X}_n(E)'_K & \xrightarrow{i^*\phi^*} & \mathrm{Hom}_A(N^{n+1}, \hat{\mathbb{G}}_{\mathrm{a}})_K & \longrightarrow & \boldsymbol{H}_{n+1}(E)_K & \longrightarrow & 0 \\
 & & \uparrow{\scriptstyle u^*} & & \uparrow{\scriptstyle u^*} & & \uparrow{\scriptstyle u^*} & & \\
0 & \longrightarrow & \boldsymbol{X}_{n-1}(E)'_K & \xrightarrow{i^*\phi^*} & \mathrm{Hom}_A(N^n, \hat{\mathbb{G}}_{\mathrm{a}})_K & \longrightarrow & \boldsymbol{H}_n(E)_K & \longrightarrow & 0 \\
 & & \uparrow & & \uparrow & & & & \\
 & & 0 & & 0 & & & &
\end{array}$$

The cokernel of each of the left two maps labeled u^* is of the displayed form by Propositions 7.4 and 8.7. If $n < m$, the expression $K\langle \phi^{\circ(n-m)*}\Theta\rangle$ is understood to be zero. The map $i^*\phi^* : K\langle \phi^{\circ(n-m)*}\Theta\rangle' \to K\langle \Psi_{n+1}\rangle$ is injective, by Proposition 8.6. Therefore the map $u^* : \boldsymbol{H}_n(E)_K \to \boldsymbol{H}_{n+1}(E)_K$ is also injective. It is an isomorphism if $n \geq m$, because $K\langle \phi^{\circ(n-m)*}\Theta\rangle$ is 1-dimensional and hence the map

$$i^*\phi^* : K\langle \phi^{\circ(n-m)*}\Theta\rangle' \to K\langle \Psi_{n+1}\rangle$$

is an isomorphism. □

Corollary 9.4. *We have*

$$\boldsymbol{H}_n(E) \otimes K \simeq \begin{cases} K\langle \Psi_1, \dots, \Psi_n\rangle & \text{if } n \leq m, \\ K\langle \Psi_1, \dots, \Psi_m\rangle & \text{if } n \geq m. \end{cases}$$

Do note that we will promote this to an integral result in Section 9B. But before we get there, we will need some preparation.

Proposition 9.5. *We have*

$$\boldsymbol{I}_n(E) \otimes K \simeq \begin{cases} K\langle \Psi_1, \dots, \Psi_n\rangle & \text{if } n \leq m-1, \\ K\langle \Psi_1, \dots, \Psi_{m-1}\rangle & \text{if } n \geq m-1. \end{cases}$$

Proof. The case $n \le m-1$ is clear. So suppose $n \ge m-1$. Then $\mathrm{Hom}_A(N^j, \hat{\mathbb{G}}_a) \otimes K$ has basis $\Psi_1, \dots, \Psi_j$, and $\boldsymbol{X}_n(E) \otimes K$ has basis $\Theta_m, \dots, (\phi^{n-m})^*\Theta_m$. Since each $(\phi^j)^*\Theta_m$ equals Ψ_{m+j} plus lower order terms, $K\langle \Psi_1, \dots, \Psi_{m-1}\rangle$ is a complement to the subspace $\boldsymbol{X}_n(E)$ of $\mathrm{Hom}_A(N^n, \hat{\mathbb{G}}_a)$. Therefore the map ∂ from $K\langle \Psi_1, \dots, \Psi_{m-1}\rangle$ to the quotient $\boldsymbol{I}_n(E)$ is an isomorphism. □

Finally the morphism $\mathrm{Hom}_A(N^n, \hat{\mathbb{G}}_a) \to \mathrm{Ext}^\sharp(E, \hat{\mathbb{G}}_a)$ of diagram (9-3) vanishes on $\phi^*(\boldsymbol{X}_{n-1}(E))$, by Proposition 9.2, and hence induces a morphism of exact sequences

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \dfrac{\boldsymbol{X}_n(E)}{\phi^*(\boldsymbol{X}_{n-1}(E)')} & \longrightarrow & \boldsymbol{H}_n(E) & \longrightarrow & \boldsymbol{I}_n(E) & \longrightarrow & 0 \\ & & \Big\downarrow \Upsilon & & \Big\downarrow \Phi & & \Big\downarrow & & \\ 0 & \longrightarrow & \mathrm{Lie}(E)^* & \longrightarrow & \mathrm{Ext}^\sharp(E, \hat{\mathbb{G}}_a) & \longrightarrow & \mathrm{Ext}(E, \hat{\mathbb{G}}_a) & \longrightarrow & 0 \end{array} \tag{9-7}$$

where as in (9-5), $\boldsymbol{I}_n(E)$ denotes the image of $\partial : \mathrm{Hom}(N^n, \hat{\mathbb{G}}_a) \to \mathrm{Ext}_A(E, \hat{\mathbb{G}}_a)$.

Proposition 9.6. *The map $\Phi : \boldsymbol{H}_n(E) \otimes K \to \mathrm{Ext}^\sharp(E, \hat{\mathbb{G}}_a) \otimes K$ is injective if and only if $\gamma \neq 0$, where $\gamma \in R$ is defined as in Proposition 8.5.*

Proof. It is enough to show that Υ is injective if and only if $\gamma \neq 0$. By Proposition 8.7, the class of Θ_m is a K-linear basis for $(\boldsymbol{X}_n(E)/\phi^*(\boldsymbol{X}_{n-1}(E)')) \otimes K$, and so it is enough to show Φ is injective if and only if $\Upsilon(\Theta_m) \neq 0$. As in (8-8), write $\Theta_m = \Psi_{\Theta_m} + g_{\Theta_m}$. Then by Proposition 9.1, it is enough to show $g'_{\Theta_m}(0) \neq 0$ if and only if $\gamma \neq 0$. But this holds because by Proposition 8.5, we have $\gamma = \pi g'_{\Theta_m}(0)$. □

Lemma 9.7. *Consider the ϕ-linear endomorphism F of K^m with matrix*

$$\begin{pmatrix} 0 & 0 & \dots & 0 & \mu_m \\ 1 & 0 & & 0 & \mu_{m-1} \\ 0 & 1 & & 0 & \mu_{m-2} \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & & 1 & \mu_1 \end{pmatrix},$$

for some given $\mu_1, \dots, \mu_m \in K$. If K^m admits an R-lattice which is stable under F, then we have $\mu_1, \dots, \mu_m \in R$.

Proof. We use Dieudonné–Manin theory. Without loss of generality, we may assume that $R/\pi R$ is algebraically closed. Let P denote the polynomial $F^m - \mu_1 F^{m-1} - \cdots - \mu_m$ in the twisted polynomial ring $K\{F\}$. Then by [Laumon 1996, B.1.5, p. 257], there exists an integer $r \ge 1$ and elements $\beta_1, \dots, \beta_m \in K(\pi^{1/r})$ such that we have

$$P = (F - \beta_1) \cdots (F - \beta_m)$$

in the ring $K(\pi^{1/r})\{F\}$ with commutation law $F\pi^{1/r} = \pi^{1/r}F$. (Note that the results of [Laumon 1996] are stated under the assumption that the residue field of R is an algebraic closure of $\mathbb{F}_p$, but they hold if it is any algebraically closed field of characteristic p.) Since $R = K \cap R[\pi^{1/r}]$, it is enough to show

$\mu_i \in R[\pi^{1/r}]$. Therefore, by replacing $R[\pi^{1/r}]$ with R, it is enough to assume that P factors as above where in addition all β_i lie in K.

Now fix i, and let us show $\beta_i \in R$. Assume $\beta_i \neq 0$, the case $\beta_i = 0$ being immediate. Because the (left) $K\{F\}$-module K^m has an F-stable integral lattice M, every quotient of K^m also has a F-stable integral lattice, namely the image of M. By [Laumon 1996, B.1.9, p. 360], for each i, the $K\{F\}$-module K^m has a quotient (in fact, a summand) isomorphic to $N = K\{F\}/K\{F\}(F - \pi^{v(\beta_i)})$. Therefore N also has a F-stable integral lattice. But this can happen only if $v(\beta_i) \geq 0$, because F sends the basis element $1 \in N$ to $\pi^{v(\beta_i)} \in N$. □

Theorem 9.8. *If E splits at m, then we have $\lambda_1, \ldots, \lambda_{m-1} \in R$, where the λ_i are as defined in Section 8C.*

Proof. We will prove the cases when $\gamma = 0$ and $\gamma \neq 0$ separately, where γ is defined as in Proposition 8.5.

Case $\gamma = 0$: When $\gamma = 0$ we have $\mathfrak{f}^* i^* = i^* \phi^*$, and hence for all $n \geq 1$, this induces a ϕ-linear map $\mathfrak{f}^* : \boldsymbol{I}_{n-1}(E) \to \boldsymbol{I}_n(E)$ as follows

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \boldsymbol{X}_n(E) & \xrightarrow{i^*} & \mathrm{Hom}_A(N^n, \hat{\mathbb{G}}_{\mathrm{a}}) & \xrightarrow{\partial} & \boldsymbol{I}_n(E) & \longrightarrow & 0 \\ & & \uparrow{\scriptstyle \phi} & & \uparrow{\scriptstyle \mathfrak{f}^*} & & \uparrow{\scriptstyle \mathfrak{f}^*} & & \\ 0 & \longrightarrow & \boldsymbol{X}_{n-1}(E) & \xrightarrow{i^*} & \mathrm{Hom}_A(N^{n-1}, \hat{\mathbb{G}}_{\mathrm{a}}) & \xrightarrow{\partial} & \boldsymbol{I}_{n-1}(E) & \longrightarrow & 0 \end{array}$$

Let $\boldsymbol{I}(E) = \varinjlim \boldsymbol{I}_n(E) \subseteq \mathrm{Ext}(E, \hat{\mathbb{G}}_{\mathrm{a}})$. Then by Proposition 9.5, the vector space $\boldsymbol{I}(E)_K$ has a K-basis $\partial\Psi_1, \ldots, \partial\Psi_{m-1}$, and with respect to this basis, the ϕ-linear endomorphism $\mathfrak{f}^*$ has matrix

$$\Gamma_0 = \begin{pmatrix} 0 & 0 & \ldots & 0 & \lambda_1 \\ 1 & 0 & & 0 & \lambda_2 \\ 0 & 1 & & 0 & \lambda_3 \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & & 1 & \lambda_{m-1} \end{pmatrix}$$

Since $\boldsymbol{I}(E)$ is contained in $\mathrm{Ext}(E, \hat{\mathbb{G}}_{\mathrm{a}})$, it is a finitely generated free R-module and hence an integral lattice in $\boldsymbol{I}(E)_K$. But then Lemma 9.7 implies $\lambda_1, \ldots, \lambda_{m-1} \in R$.

Case $\gamma \neq 0$: Let $\boldsymbol{H}(E) = \varinjlim \boldsymbol{H}_n(E)$. Let us consider the matrix Γ of the ϕ-linear endomorphism $\mathfrak{f}$ of $\boldsymbol{H}(E)_K$ with respect to the K-basis $\Psi_1, \ldots, \Psi_m$ given by Corollary 9.4. Then by Proposition 8.5 and (8-12), we have

$$\begin{aligned} i^* \phi^* \Theta_m &= \mathfrak{f}^* i^* \Theta_m + \gamma \Psi_1 \\ &= \mathfrak{f}^*(\Psi_m - \lambda_{m-1}\Psi_{m-1} - \cdots - \lambda_1 \Psi_1) + \gamma \Psi_1 \\ &= \mathfrak{f}^*(\Psi_m) - \phi(\lambda_{m-1})\Psi_m - \cdots - \phi(\lambda_1)\Psi_2 + \gamma \Psi_1. \end{aligned}$$

Therefore we have

$$\mathfrak{f}^*(\Psi_m) \equiv \phi(\lambda_{m-1})\Psi_m + \cdots + \phi(\lambda_1)\Psi_2 - \gamma \Psi_1 \bmod i^* \phi^*(\boldsymbol{X}'_m)$$

and hence

$$\Gamma = \begin{pmatrix} 0 & 0 & \dots & 0 & -\gamma \\ 1 & 0 & & 0 & \phi(\lambda_1) \\ 0 & 1 & & 0 & \phi(\lambda_2) \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & & 0 & \phi(\lambda_{m-2}) \\ 0 & 0 & & 1 & \phi(\lambda_{m-1}) \end{pmatrix}$$

We will now apply Lemma 9.7 to the operator $\mathfrak{f}^*$ on $\boldsymbol{H}(E)_K$, but to do this we need to produce an integral lattice M. Consider the commutative square

$$\begin{array}{ccc} \boldsymbol{H}(E) & \xrightarrow{\Phi} & \mathrm{Ext}^\sharp(E, \hat{\mathbb{G}}_{\mathrm{a}}) \\ \downarrow & & \downarrow j \\ \boldsymbol{H}(E)_K & \xrightarrow{\Phi_K} & \mathrm{Ext}^\sharp(E, \hat{\mathbb{G}}_{\mathrm{a}})_K. \end{array}$$

Let M denote the image of $\boldsymbol{H}(E)$ in $\boldsymbol{H}(E)_K$. It is clearly stable under $\mathfrak{f}^*$. But also the maps Φ_K and j are injective, by Proposition 9.6 and because $\mathrm{Ext}^\sharp(E, \hat{\mathbb{G}}_{\mathrm{a}}) \simeq R^r$; so M agrees with the image of $\boldsymbol{H}(E)$ in $\mathrm{Ext}^\sharp(E, \hat{\mathbb{G}}_{\mathrm{a}})$ and is therefore finitely generated.

We can then apply Lemma 9.7 and deduce $\phi(\lambda_{m-1}), \dots, \phi(\lambda_1) \in R$. This implies $\lambda_{m-1}, \dots, \lambda_1 \in R$, since $R/\pi R$ is a field and hence the Frobenius map on it is injective. □

Corollary 9.9. (1) *The element $\Theta_m \in \boldsymbol{X}_m(E)_K$ lies in $\boldsymbol{X}_m(E)$.*

(2) *For $n \geq m$, all the maps in the diagram*

$$\begin{array}{ccc} \boldsymbol{X}_n(E)/\boldsymbol{X}_{n-1}(E) & \xrightarrow{\phi^*} & \boldsymbol{X}_{n+1}(E)/\boldsymbol{X}_n(E) \\ \downarrow i^* & & \downarrow i^* \\ \mathrm{Hom}_A(N^n, \hat{\mathbb{G}}_{\mathrm{a}})/\mathrm{Hom}_A(N^{n-1}, \hat{\mathbb{G}}_{\mathrm{a}}) & \xrightarrow{\mathfrak{f}^*} & \mathrm{Hom}_A(N^{n+1}, \hat{\mathbb{G}}_{\mathrm{a}})/\mathrm{Hom}_A(N^n, \hat{\mathbb{G}}_{\mathrm{a}}) \end{array}$$

are isomorphisms.

Proof. (1) By Theorem 9.8, the element $i^*\Theta_m$ of $\mathrm{Hom}_A(N^m, \hat{\mathbb{G}}_{\mathrm{a}})_K$ actually lies in $\mathrm{Hom}_A(N^m, \hat{\mathbb{G}}_{\mathrm{a}})$, and therefore by the exact sequence (8-3) we have $\Theta_m \in \boldsymbol{X}_m(E)$.

(2) By Proposition 8.6, we know $\mathfrak{f}^*$ is an isomorphism.

Also by Proposition 8.6, the maps i^* are injective for all $n \geq m$. So to show they are isomorphisms, it is enough to show they are surjective. The R-linear generator Ψ_m of $\mathrm{Hom}_A(N^n, \hat{\mathbb{G}}_{\mathrm{a}})/\mathrm{Hom}_A(N^{n-1}, \hat{\mathbb{G}}_{\mathrm{a}})$ is the image of Θ_m, which by part (1), lies in $\boldsymbol{X}_m(E)$. Therefore i^* is surjective for $n = m$. Then because $\mathfrak{f}^*$ is an isomorphism, it follows by induction that i^* is surjective for all $n \geq m$.

Finally, ϕ^* is an isomorphism because all the other morphisms in the diagram are. □

We knew before that $i^*(\phi^j)^*\Theta_m$ agrees with Ψ_{m+j} plus lower order rational characters, but the corollary above implies that these lower order characters are in fact integral.

Theorem 9.10. *Let E be a Drinfeld module that splits at m.*

(1) *For any $n \geq m$, the composition*

$$\boldsymbol{X}_n(E) \to \mathrm{Hom}_A(N^n, \hat{\mathbb{G}}_{\mathrm{a}}) \to \mathrm{Hom}_A(N^n, \hat{\mathbb{G}}_{\mathrm{a}})/\mathrm{Hom}_A(N^{m-1}, \hat{\mathbb{G}}_{\mathrm{a}}) \tag{9-8}$$

is an isomorphism of R-modules.

(2) *$\boldsymbol{X}_n(E)$ is freely generated as an R-module by $\Theta_m, \dots, (\phi^*)^{n-m}\Theta_m$.*

Proof. (1) By Corollary 9.9, the induced morphism on each graded piece is an isomorphism. It follows that the map in question is also an isomorphism.

(2) This follows formally from (1) and the fact, which follows from Corollary 9.9, that the map (9-8) sends any $(\phi^*)^j\Theta_m$ to Ψ_{m+j} plus lower order terms. □

9B. *$\boldsymbol{H}(E)$ and de Rham cohomology.* Collecting the results above, we can now prove Theorem 1.2. Let m denote the splitting order of E, as defined in section 8. We have isomorphisms

$$R\langle\Psi_1, \dots, \Psi_{m-1}\rangle = \mathrm{Hom}_A(N^{m-1}, \hat{\mathbb{G}}_{\mathrm{a}}) \xrightarrow{\sim} \boldsymbol{I}_n(E) \quad \text{and} \quad R\langle\Psi_1, \dots, \Psi_m\rangle = \mathrm{Hom}_A(N^m, \hat{\mathbb{G}}_{\mathrm{a}}) \xrightarrow{\sim} \boldsymbol{H}_n(E)$$

for $n \geq m$, and hence in the limit

$$R\langle\Psi_1, \dots, \Psi_{m-1}\rangle \xrightarrow{\sim} \boldsymbol{I}(E) \quad \text{and} \quad R\langle\Psi_1, \dots, \Psi_m\rangle \xrightarrow{\sim} \boldsymbol{H}(E)$$

And so the K-linear bases of $K \otimes \boldsymbol{I}(E)$ and $K \otimes \boldsymbol{H}(E)$ — the ones respect to which the action of $\mathfrak{f}^*$ is described by the matrices Γ_0 and Γ in the proof of Theorem 9.8 — are in fact R-linear bases of $\boldsymbol{I}(E)$ and $\boldsymbol{H}(E)$.

We also have isomorphisms for $n \geq m$

$$R\langle\Theta_m\rangle = \boldsymbol{X}_m(E) \xrightarrow{\sim} \boldsymbol{X}_n(E)/\phi^*(\boldsymbol{X}_{n-1}(E)').$$

Combining these, we have the following map between exact sequences of R-modules, as in (9-7):

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \boldsymbol{X}_m(E) & \longrightarrow & \boldsymbol{H}(E) & \longrightarrow & \boldsymbol{I}(E) & \longrightarrow & 0 \\ & & \Big\downarrow{\scriptstyle \Upsilon} & & \Big\downarrow{\scriptstyle \Phi} & & \Big\downarrow & & \\ 0 & \longrightarrow & \mathrm{Lie}(E)^* & \longrightarrow & \mathrm{Ext}^\sharp(E, \hat{\mathbb{G}}_{\mathrm{a}}) & \longrightarrow & \mathrm{Ext}(E, \hat{\mathbb{G}}_{\mathrm{a}}) & \longrightarrow & 0 \end{array}$$

where Υ sends Θ_m to γ/π (in coordinates). It follows that Φ is injective if and only if $\gamma \neq 0$.

10. Computation of λ_1 and γ in the rank 2 case

In this section, we compute λ_1 and γ for Drinfeld modules of rank 2, the first nontrivial case. Recall from (8-8), Proposition 8.5(1), and (8-12) that we have

$$\Theta_2 = \Psi_2(x', x'') - \lambda_1\Psi_1(x') + \pi^{-1}\gamma x + (\text{higher-degree terms in } x) \tag{10-1}$$

assuming the splitting number m is 2. The result below shows that λ_1 and γ depend on the higher Buium derivatives $a_i', a_i'', \ldots$ of the modular parameters a_i, and not only on the modular parameters themselves. So it seems that our F-crystal $\boldsymbol{H}$ is not determined by the classical realizations, such as the crystalline realization or the Tate module, in any straightforward manner.

Theorem 10.1. *Let $A = \mathbb{F}_q[v]$ with $q \geq 3$, let $t \in A$ be an irreducible polynomial of degree f, and let E be a Drinfeld module over R satisfying*

$$\varphi_E(t)(x) = \pi x + a_1 x^q + a_2 x^{q^2}. \tag{10-2}$$

Then we have

$$\lambda_1 \equiv (-1)^f w^{(q^{f-1}(q^f-1))/(q-1)}(1 - a_1' w^{q^{f-1}} + a_2' w^{q^{f-1}+q^f})^{q^f-1} \mod \pi,$$

where $w = a_1 a_2^{-1}$, and

$$\gamma \mod \pi^2 \equiv \begin{cases} \pi\lambda_1/a_1 & \text{if } f = 1, \\ -\pi\lambda_1/a_2 & \text{if } a_1 \equiv 0 \bmod \pi \text{ and } f = 2, \\ 0 & \text{if } a_1 \not\equiv 0 \bmod \pi \text{ or } f \geq 3. \end{cases}$$

Observe that when $\varphi_E(t)(x)$ is of the form $\pi x + ax^q + x^{q^2}$, which is always true after changing the coordinate x (perhaps passing to a cover of S), we have the simplified forms

$$\lambda_1 \equiv (-1)^f a^{(q^{f-1}(q^f-1))/(q-1)}(1 - a' a^{q^{f-1}})^{q^f-1} \mod \pi, \tag{10-3}$$

$$\gamma \bmod \pi^2 \equiv \begin{cases} \pi\lambda_1/a & \text{if } f = 1, \\ -\pi\lambda_1 & \text{if } a \equiv 0 \bmod \pi \text{ and } f = 2, \\ 0 & \text{if } a \not\equiv 0 \bmod \pi \text{ or } f \geq 3. \end{cases} \tag{10-4}$$

Proof. Let $\vartheta_1 : N^1 \to \hat{\mathbb{G}}_a$ be the isomorphism defined in Theorem 6.2. Then $\vartheta_1 \equiv \tau^0 \bmod \pi$. Also ϑ_1 induces the isomorphism $(\vartheta_1)_* : \mathrm{Ext}(E, N^1) \to \mathrm{Ext}(E, \hat{\mathbb{G}}_a)$. In order to determine the action of A on J^1E and J^2E we need to determine how t acts on the coordinates x' and x''. Now we note that $J^nE \simeq W_n$ can be endowed with the δ-coordinates (denoted $[z, z', z'', \ldots]$) or the Witt coordinates (denoted $(z_0, z_1, z_2, \ldots)$) and they are related by the following in J^2E by Proposition 3.2

$$[z, z', z''] = (z, z', z'' + \pi^{\hat{q}-2}(z')^{\hat{q}}). \tag{10-5}$$

Taking π-derivatives of both sides of (10-2) using the formula

$$\delta(ax^{q^j}) = a' x^{\hat{q}q^i} + \phi(a)\pi^{q^i-1}(x')^{q^i},$$

we obtain

$$\varphi(t)(x') = \pi' x^{\hat{q}} + a_1' x^{q\hat{q}} + a_2' x^{q^2\hat{q}} + \pi x' + \phi(a_1)\pi^{q-1}(x')^q + \phi(a_2)\pi^{q^2-1}(x')^{q^2} \tag{10-6}$$

and

$$\varphi(t)(x'') = \pi'' x^{\hat{q}^2} + a_1'' x^{q\hat{q}^2} + a_2'' x^{q^2\hat{q}^2} + \{\text{terms with } x' \text{ and } x''\}. \tag{10-7}$$

Then the A-action $\varphi_{J^1E} : A \to \mathrm{End}(J^1E)$ is given in Witt coordinates by the 2×2 matrix

$$\varphi_{J^1E}(t) = \begin{pmatrix} \varphi_E(t) & 0 \\ \eta_{J^1E} & \varphi_{N^1}(t) \end{pmatrix}$$

where $\eta_{J^1E} = \pi' x^{\hat{q}} + a_1' x^{q\hat{q}} + a_2' x^{q^2\hat{q}}$. By (10-7) and (10-5), the A-action $A \to \mathrm{End}(J^2E)$ is given by the $(1+2) \times (1+2)$ block matrix

$$\varphi_{J^2E}(t) = \begin{pmatrix} \varphi_E(t) & 0 \\ \eta_{J^2E} & \varphi_{N^2}(t) \end{pmatrix}$$

where (using (10-5)) η_{J^2E} is the column vector

$$\eta_{J^2E} = \begin{pmatrix} \pi' x^{\hat{q}} + a_1' x^{q\hat{q}} + a_2' x^{q^2\hat{q}} \\ \Delta(\pi) x^{\hat{q}^2} + \Delta(a_1) x^{q\hat{q}^2} + \Delta(a_2) x^{q^2\hat{q}^2} \end{pmatrix}$$

and where $\Delta(y) = y'' + \pi^{\hat{q}-2}(y')^{\hat{q}}$.

Now we will consider two cases:

(1) Consider $\eta_{\Psi_{1*}(J^1E)} \in \mathrm{Ext}(E, \hat{\mathbb{G}}_{\mathrm{a}})$ which is the image of Ψ_1 under the connecting morphism

$$\mathrm{Hom}_A(\hat{\mathbb{G}}_{\mathrm{a}}, \hat{\mathbb{G}}_{\mathrm{a}}) \xrightarrow{\partial} \mathrm{Ext}(E, \hat{\mathbb{G}}_{\mathrm{a}})$$

and $\Psi_1 = \vartheta_1 : N^1 \to \hat{\mathbb{G}}_{\mathrm{a}}$ is the isomorphism defined in Theorem 6.2 and satisfies $\Psi_1 = \tau^0 \bmod \pi$.

$$\begin{array}{ccccccccc} 0 & \longrightarrow & N^1 & \longrightarrow & J^1E & \longrightarrow & E & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle \Psi_1} & & \downarrow & & \| & & \\ 0 & \longrightarrow & \hat{\mathbb{G}}_{\mathrm{a}} & \longrightarrow & f_*(J^1E) & \longrightarrow & E & \longrightarrow & 0 \end{array}$$

where $\eta_{J^1E} = [\pi' x^{\hat{q}} + a_1' x^{q\hat{q}} + a_2' x^{q^2\hat{q}}] \in \mathrm{Ext}(E, N^1)$. Hence

$$\eta_{\Psi_{1*}(J^1E)} = [\pi' x^{\hat{q}} + a_1' x^{q\hat{q}} + a_2' x^{q^2\hat{q}}] \in \mathrm{Ext}(E, \hat{\mathbb{G}}_{\mathrm{a}}) \quad \text{and} \quad \partial(\Psi_1) \equiv [x^{\hat{q}} + a_1' x^{q\hat{q}} + a_2' x^{q^2\hat{q}}] \bmod \pi.$$

(2) Now consider $\eta_{\Psi_{2*}(J^2E)} \in \mathrm{Ext}(E, \hat{\mathbb{G}}_{\mathrm{a}})$ obtained as

$$\begin{array}{ccccccccc} 0 & \longrightarrow & N^2 & \longrightarrow & J^2E & \longrightarrow & E & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle \Psi_2} & & \downarrow & & \| & & \\ 0 & \longrightarrow & \hat{\mathbb{G}}_{\mathrm{a}} & \longrightarrow & f_*(J^2E) & \longrightarrow & E & \longrightarrow & 0 \end{array}$$

Now we have

$$\eta_{J^2E} = \left[\begin{pmatrix} \pi' x^{\hat{q}} + a_1' x^{q\hat{q}} + a_2' x^{q^2\hat{q}} \\ \Delta(\pi) x^{\hat{q}^2} + \Delta(a_1) x^{q\hat{q}^2} + \Delta(a_2) x^{q^2\hat{q}^2} \end{pmatrix} \right] \in \mathrm{Ext}(E, N^2)$$

Let $\mathcal{F}(y) = (y')^{\hat{q}} + \pi\Delta(y)$. Then applying $\Psi_2 = \vartheta_1 \circ \mathfrak{f}$ and $\mathfrak{f}(z_1, z_2) = z_1^{\hat{q}} + \pi z_2$, we have

$$\begin{aligned}
\partial(\Psi_2) = \eta_{\Psi_{2*}(J^2E)} &= [\vartheta_1(\mathcal{F}(\pi)x^{\hat{q}^2} + \mathcal{F}(a_1)x^{q\hat{q}^2} + \mathcal{F}(a_2)x^{q^2\hat{q}^2})] \in \mathrm{Ext}(E, \hat{\mathbb{G}}_{\mathrm{a}}) \\
\partial(\Psi_2) &\equiv [\mathcal{F}(\pi)x^{\hat{q}^2} + \mathcal{F}(a_1)x^{q\hat{q}^2} + \mathcal{F}(a_2)x^{q^2\hat{q}^2}] \bmod \pi \\
&\equiv [(\pi')^{\hat{q}}x^{\hat{q}^2} + (a_1')^{\hat{q}}x^{q\hat{q}^2} + (a_2')^{\hat{q}}x^{q^2\hat{q}^2}] \bmod \pi \\
&\equiv [x^{\hat{q}^2} + (a_1')^{\hat{q}}x^{q\hat{q}^2} + (a_2')^{\hat{q}}x^{q^2\hat{q}^2}] \bmod \pi.
\end{aligned}$$

Recall [Gekeler 1990a, §5] that the map $R\{\tau\}\hat{} \to \mathrm{Ext}(E, \hat{\mathbb{G}}_{\mathrm{a}})$ given by $\eta \mapsto [\eta]$ is surjective and the kernel consists of the inner derivations, which is to say all η of the form

$$\pi\alpha - \alpha \circ \varphi_E(t),$$

for some $\alpha \in R\{\tau\}\hat{}$. Let us now work out these relations explicitly for $\alpha = \tau^0, \tau^1, \tau^2$. If $\alpha = \tau^j$, with $j \geq 0$, we get the relation

$$\begin{aligned}
\pi\tau^j &= \tau^j(\pi\tau^0 + a_1\tau^1 + a_2\tau^2) \\
\tau^{j+2} &= a_2^{-q^j}[(\pi - \pi^{q^j})\tau^j - a_1^{q^j}\tau^{j+1}] \\
\tau^{j+2} &\equiv -(a_1a_2^{-1})^{q^j}\tau^{j+1} \bmod \pi
\end{aligned}$$

and hence we have by induction the relations

$$\tau^{i+1} \equiv (-1)^i w^{(q^i-1)/(q-1)}\tau^1 \bmod \pi \tag{10-8}$$

where $w = a_1a_2^{-1}$, for all $i \geq 0$.

Therefore writing $\hat{q} = q^f$, we have

$$\begin{aligned}
\partial(\Psi_1) &\equiv x^{\hat{q}} + a_1'x^{q\hat{q}} + a_2'x^{q^2\hat{q}} \\
&\equiv x^{q^f} + a_1'x^{q^{f+1}} + a_2'x^{q^{f+2}} \\
&\equiv \tau^f + a_1'\tau^{f+1} + a_2'\tau^{f+2} \\
&\equiv (-1)^{f+1}w^{1+\cdots+q^{f-2}}(1 - a_1'w^{q^{f-1}} + a_2'w^{q^{f-1}+q^f})\tau^1
\end{aligned}$$

and

$$\begin{aligned}
\partial(\Psi_2) &\equiv x^{\hat{q}^2} + (a_1')^{\hat{q}}x^{q\hat{q}^2} + (a_2')^{\hat{q}}x^{q^2\hat{q}^2} \\
&\equiv \tau^{2f} + (a_1')^{q^f}\tau^{2f+1} + (a_2')^{q^f}\tau^{2f+2} \\
&\equiv (-1)^{2f+1}w^{1+\cdots+q^{2f-2}}(1 - (a_1')^{q^f}w^{q^{2f-1}} + (a_2')^{q^f}w^{q^{2f-1}+q^{2f}})\tau^1 \\
&\equiv (-1)^{2f+1}w^{1+\cdots+q^{2f-2}}(1 - a_1'w^{q^{f-1}} + a_2'w^{q^{f-1}+q^f})^{q^f}\tau^1.
\end{aligned}$$

and hence

$$\begin{aligned}
\lambda_1 = \frac{\partial(\Psi_2)}{\partial(\Psi_1)} &\equiv (-1)^f w^{q^{f-1}+\cdots+q^{2f-2}}(1 - a_1'w^{q^{f-1}} + a_2'w^{q^{f-1}+q^f})^{q^f-1} \bmod \pi \\
&\equiv (-1)^f w^{q^{f-1}(1+\cdots+q^{f-1})}(1 - a_1'w^{q^{f-1}} + a_2'w^{q^{f-1}+q^f})^{q^f-1} \bmod \pi \\
&\equiv (-1)^f w^{(q^{f-1}(q^f-1))/(q-1)}(1 - a_1'w^{q^{f-1}} + a_2'w^{q^{f-1}+q^f})^{q^f-1} \bmod \pi
\end{aligned}$$

Now we determine γ. Write $g = g_{\Theta_2} = \sum_i \alpha_i \tau^i$. Then from Proposition 8.5, we know $\gamma = \pi\alpha_0$. Now we will compute α_0. Let $(z_0, z_1, z_2) := \varphi_{J^2E}(t)(x, 0, 0)$. Then

$$\begin{aligned}\Theta_2(\varphi_{J^2E}(t)(x,0,0)) &= \Psi_2(z_1,z_2) - \lambda_1\Psi_1(z_1) + g(z_0)\\ &= \vartheta_1(z_1^{\hat{q}} + \pi z_2) - \lambda_1\vartheta_1(z_1) + g(z_0)\\ &\equiv z_1^{\hat{q}} - \lambda_1 z_1 + g(z_0) \bmod \pi\end{aligned}$$

where $z_0 = \pi x + a_1x^q + a_2x^{q^2}$ and $z_1 = \pi' x^{\hat{q}} + a_1' x^{q\hat{q}} + a_2' x^{q^2\hat{q}}$. On the other hand from the A-linearity of Θ_2 we have

$$\Theta_2(\varphi_{J^2E}(t)(x,0,0)) = \varphi_{\hat{\mathbb{G}}_a}(t)\Theta_2(x,0,0) = \pi\Theta_2(x,0,0) \equiv 0 \bmod \pi$$

and hence $z_1^{\hat{q}} - \lambda_1 z_1 + g(z_0) \equiv 0 \bmod \pi$. Substituting z_0 and z_1 in, we obtain

$$\begin{aligned}0 &\equiv (\pi' x^{\hat{q}} + a_1' x^{q\hat{q}} + a_2' x^{q^2\hat{q}})^{\hat{q}} - \lambda_1(\pi' x^{\hat{q}} + a_1' x^{q\hat{q}} + a_2' x^{q^2\hat{q}}) + g(\pi x + a_1x^q + a_2x^{q^2})\\ &\equiv (x^{\hat{q}} + a_1' x^{q\hat{q}} + a_2' x^{q^2\hat{q}})^{\hat{q}} - \lambda_1(x^{\hat{q}} + a_1' x^{q\hat{q}} + a_2' x^{q^2\hat{q}}) + g(a_1x^q + a_2x^{q^2})\end{aligned}$$

Now substitute $g(x) = \sum_{j\geq 0}\alpha_j x^{q^j}$ into this and consider the coefficient of x^q. If $\hat{q} = q$, we obtain $\lambda_1 \equiv \alpha_0 a_1$ and hence

$$\gamma = \pi\alpha_0 \equiv \frac{\pi\lambda_1}{a_1} \bmod \pi^2.$$

If $\hat{q} \neq q$, we obtain $\alpha_0 a_1 \equiv 0$ and hence $\gamma \equiv 0 \bmod \pi^2$ if $a_1 \not\equiv 0 \mod \pi$. If $a_1 \equiv 0 \mod \pi$, we consider the coefficient of x^{q^2} which is $\alpha_0 a_2 + \lambda_1$ if $f = 2$ and $\alpha_0 a_2$ otherwise. In the case when $f = 2$ we have $\alpha_0 \equiv \lambda_1/a_2 \mod \pi$ since a_2 is invertible and hence $\gamma \equiv -\pi\lambda_1/a_2 \mod \pi^2$. When $f \geq 3$ we have $\alpha_0 \equiv 0 \mod \pi$ and hence the result follows. □

Acknowledgement.

We wish to thank the anonymous referee for carefully reading our article and the suggestions which led to deeper clarifications and brought more lucidity in our present version of the paper.

References

[Borger 2011a] J. Borger, "The basic geometry of Witt vectors, I: The affine case", *Algebra Number Theory* **5**:2 (2011), 231–285. MR Zbl

[Borger 2011b] J. Borger, "The basic geometry of Witt vectors, II: Spaces", *Math. Ann.* **351**:4 (2011), 877–933. MR Zbl

[Borger and Saha 2017a] J. Borger and A. Saha, "Isocrystals associated to arithmetic jet spaces of abelian schemes", preprint, 2017. To appear in *Advances in Mathematics*. arXiv

[Borger and Saha 2017b] J. Borger and A. Saha, "On Frobenius and fibers of arithmetic jet spaces", preprint, 2017. arXiv

[Buium 1992] A. Buium, "Intersections in jet spaces and a conjecture of S. Lang", *Ann. of Math.* (2) **136**:3 (1992), 557–567. MR Zbl

[Buium 1995] A. Buium, "Differential characters of abelian varieties over p-adic fields", *Invent. Math.* **122**:2 (1995), 309–340. MR Zbl

[Buium 2000] A. Buium, "Differential modular forms", *J. Reine Angew. Math.* **520** (2000), 95–167. MR Zbl

[Buium and Saha 2011] A. Buium and A. Saha, "Differential overconvergence", pp. 99–129 in *Algebraic methods in dynamical systems*, edited by T. Crespo and Z. Hajto, Banach Center Publ. **94**, Polish Acad. Sci. Inst. Math., Warsaw, 2011. MR Zbl

[Buium and Saha 2012a] A. Buium and A. Saha, "Hecke operators on differential modular forms mod p", *J. Number Theory* **132**:5 (2012), 966–997. MR Zbl

[Buium and Saha 2012b] A. Buium and A. Saha, "The ring of differential Fourier expansions", *J. Number Theory* **132**:5 (2012), 896–937. MR Zbl

[Buium and Saha 2014] A. Buium and A. Saha, "The first p-jet space of an elliptic curve: global functions and lifts of Frobenius", *Math. Res. Lett.* **21**:4 (2014), 677–689. MR Zbl

[Drinfeld 1974] V. G. Drinfeld, "Elliptic modules", *Mat. Sb.* (*N.S.*) **94(136)** (1974), 594–627, 656. In Russian; translated in *Math. USSR-Sb.* **23**:4 (1974), 561–592. MR Zbl

[Drinfeld 1976] V. G. Drinfeld, "Coverings of p-adic symmetric domains", *Funkcional. Anal. i Priložen.* **10**:2 (1976), 29–40. In Russian; translated in *Funct. Anal. Appl.* **10**:2 (1976), 107–115. MR Zbl

[Dupuy 2014] T. Dupuy, "Deligne–Illusie classes, I: Lifted torsors of lifts of the Frobenius for curves", preprint, 2014. arXiv

[Gekeler 1990a] E.-U. Gekeler, "de Rham cohomology and the Gauss–Manin connection for Drinfeld modules", pp. 223–255 in *p-adic analysis* (Trento, 1989), edited by F. Baldassarri et al., Lecture Notes in Math. **1454**, Springer, 1990. MR Zbl

[Gekeler 1990b] E.-U. Gekeler, "de Rham cohomology for Drinfeld modules", pp. 57–85 in *Séminaire de théorie des nombres, Paris 1988–1989*, edited by C. Goldstein, Progr. Math. **91**, Birkhäuser, Boston, 1990. MR Zbl

[Goss 1996] D. Goss, *Basic structures of function field arithmetic*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)] **35**, Springer, 1996. MR Zbl

[Hartl 2017] U. Hartl, "Isogenies of abelian Anderson A-modules and A-motives", preprint, 2017. arXiv

[Hesselholt 2015] L. Hesselholt, "The big de Rham–Witt complex", *Acta Math.* **214**:1 (2015), 135–207. MR Zbl

[Joyal 1985] A. Joyal, "δ-anneaux et vecteurs de Witt", *C. R. Math. Rep. Acad. Sci. Canada* **7**:3 (1985), 177–182. MR Zbl

[Laumon 1996] G. Laumon, *Cohomology of Drinfeld modular varieties, Part I*, Cambridge Studies in Advanced Mathematics **41**, Cambridge University Press, 1996. MR Zbl

[Manin 1963] J. I. Manin, "Rational points on algebraic curves over function fields", *Izv. Akad. Nauk SSSR Ser. Mat.* **27** (1963), 1395–1440. Translated as "Rational points of algebraic curves over function fields", pp. 189–234 in *Fifteen papers on algebra*, Transl. Amer. Math. Soc. (2) **50**, Amer. Math. Soc., Providence, RI, 1966. MR Zbl

[Witt 1937] E. Witt, "Zyklische Körper und Algebren der Charakteristik p vom Grad p^n. Struktur diskret bewerteter perfekter Körper mit vollkommenem Restklassenkörper der Charakteristik p", *J. Reine Angew. Math.* **176** (1937), 126–140. MR

Communicated by Kiran S. Kedlaya
Received 2017-09-03 Revised 2018-11-26 Accepted 2019-02-22

james.borger@anu.edu.au *Mathematical Sciences Institute, Australian National University, Canberra ACT, Australia*

arnabsaha0930@gmail.com *Max Planck Institute for Mathematics, Bonn, Germany*

msp

ALGEBRA AND NUMBER THEORY 13:4 (2019)
dx.doi.org/10.2140/ant.2019.13.839

Quadratic twists of abelian varieties and disparity in Selmer ranks

Adam Morgan

We study the parity of 2-Selmer ranks in the family of quadratic twists of a fixed principally polarized abelian variety over a number field. Specifically, we determine the proportion of twists having odd (respectively even) 2-Selmer rank. This generalizes work of Klagsbrun–Mazur–Rubin for elliptic curves and Yu for Jacobians of hyperelliptic curves. Several differences in the statistics arise due to the possibility that the Shafarevich–Tate group (if finite) may have order twice a square. In particular, the statistics for parities of 2-Selmer ranks and 2-infinity Selmer ranks need no longer agree and we describe both.

1. Introduction

In this paper we study how various invariants of principally polarized abelian varieties behave under quadratic twist.

Our first result determines the distribution of the parities of 2-Selmer ranks in the quadratic twist family of an arbitrary principally polarized abelian variety. Specifically, for a number field K (with absolute Galois group G_K) and real number $X > 0$ set

$$\mathcal{C}(K, X) = \{\chi \in \mathrm{Hom}_{\mathrm{cnt}}(G_K, \{\pm 1\}) : \mathrm{Norm}(\mathfrak{p}) < X \text{ for all primes } \mathfrak{p} \text{ at which } \chi \text{ ramifies}\}.$$

MSC2010: 11G10.
Keywords: Abelian varieties, Selmer groups, quadratic twist, ranks, Shafarevich–Tate group.

Theorem 1.1. *Let A/K be a principally polarized abelian variety and let*

$$\epsilon : \mathrm{Gal}(K(A[2])/K) \to \{\pm 1\}$$

be the map

$$\sigma \mapsto (-1)^{\dim_{\mathbb{F}_2} A[2]^\sigma}.$$

(i) *If ϵ is a homomorphism then, for all sufficiently large X,*

$$\frac{|\{\chi \in \mathcal{C}(K, X) : \dim_{\mathbb{F}_2} \mathrm{Sel}_2(A^\chi/K) \text{ is even}\}|}{|\mathcal{C}(K, X)|} = \frac{1 + (-1)^{\dim_{\mathbb{F}_2} \mathrm{Sel}_2(A/K)} \cdot \delta}{2}$$

where δ is a finite product of explicit local terms δ_v (see the statement of Theorem 10.13 for their definition).

(ii) *If ϵ fails to be a homomorphism then, for all sufficiently large X,*

$$\frac{|\{\chi \in \mathcal{C}(K, X) : \dim_{\mathbb{F}_2} \mathrm{Sel}_2(A^\chi/K) \text{ is even}\}|}{|\mathcal{C}(K, X)|} = \frac{1}{2}.$$

(*Here for $\chi \in \mathrm{Hom}_{\mathrm{cnt}}(\mathrm{Gal}(\overline{K}/K), \{\pm 1\})$ we let A^χ/K denote the quadratic twist of A by χ.*)

Theorem 1.1 is known for elliptic curves by work of Klagsbrun, Mazur and Rubin [2013, Theorem A] and, more generally, for Jacobians of odd degree hyperelliptic curves by work of Yu [2016, Theorem 1]. These previous results both fall into case (i) of Theorem 1.1, thus the failure of ϵ to be a homomorphism forcing parity in the distribution is a phenomenon new to this work. Despite this, case (ii) of Theorem 1.1 is in some sense the "generic" case since if $\mathrm{Gal}(K(A[2])/K)$ is the full symplectic group $\mathrm{Sp}_{2g}(\mathbb{F}_2)$ for $g = \dim A \geq 3$ then the simplicity of $\mathrm{Sp}_{2g}(\mathbb{F}_2)$ prevents ϵ from being a homomorphism. For a discussion of when ϵ is or is not a homomorphism for various families of abelian varieties, see Section 10C.

In the two previously known cases above, finiteness of the 2-primary subgroup of the Shafarevich–Tate group is known to imply that the parity of the 2-Selmer rank agrees with that of the Mordell–Weil rank, so that Theorem 1.1 is conjecturally satisfied by Mordell–Weil ranks also. For general principally polarized abelian varieties, however, this need not be true due to a phenomenon first observed by Poonen and Stoll [1999]: the 2-primary subgroup of the Shafarevich–Tate group, if finite, need not have square order. Thus to see how one expects the parity of Mordell–Weil ranks to behave in quadratic twist families we also prove a version of Theorem 1.1 for 2^∞-Selmer ranks (by definition the 2^∞-Selmer rank, denoted rk_2, is equal to the sum of the Mordell–Weil rank and the (conjecturally trivial) $\mathbb{Z}_2$-corank of the 2-primary subgroup of the Shafarevich–Tate group).

Theorem 1.2. *Let A/K be a principally polarized abelian variety. Then, for all sufficiently large $X > 0$,*

$$\frac{|\{\chi \in \mathcal{C}(K, X) : \mathrm{rk}_2(A^\chi/K) \text{ is even}\}|}{|\mathcal{C}(K, X)|} = \frac{1 + (-1)^{\mathrm{rk}_2(A/K)} \cdot \kappa}{2}$$

where κ is an explicit finite product of local terms κ_v given in Definition 10.21.

We remark that if $\dim A$ is odd and K has a real place then $\kappa = 0$. In general however κ is often nonzero: see [Klagsbrun et al. 2013, Example 7.11] for an example of an elliptic curve for which κ is dense in $[-1, 1]$ as the base field K varies, and [Yu 2016, Proposition 8.1] for an example of an abelian surface over $\mathbb{Q}$ for which $\kappa = 1$.

Combining Theorems 1.1 and 1.2 we see that the distribution of parities of 2-Selmer ranks and 2^∞-Selmer ranks in general behave quite differently, as the following example illustrates.

Example 1.3 (see Example 10.24). Let $J/\mathbb{Q}$ be the Jacobian of the genus 2 hyperelliptic curve $C : y^2 = x^6 + x^4 + x + 3$. Then the function ϵ is not a homomorphism for $J/\mathbb{Q}$ so that half of the 2-Selmer ranks of the quadratic twists of J are even and are half odd. On the other hand, J has $\kappa = \frac{3}{16}$ and odd 2^∞-Selmer rank, so that $\frac{19}{32}$ of the twists of J have even 2^∞-Selmer rank and $\frac{13}{32}$ have odd 2^∞-Selmer rank.

In fact, in the case where ϵ fails to be a homomorphism we show that the parity of the 2^∞-Selmer ranks behaves in some sense independently of the parity of the 2-Selmer ranks. See Remark 10.26 for the proof of this statement and Corollary 10.29 for a description of the joint distribution of the parities of 2-Selmer ranks and 2-infinity Selmer ranks in all cases.

A key step in passing between Theorem 1.1 and Theorem 1.2 is the study of how the "nonsquare order Shafarevich–Tate group" phenomenon behaves under quadratic twist. Our main result here is:

Theorem 1.4. *Let A/K be an abelian variety equipped with a principal polarization λ defined over K and let $\chi \in \mathrm{Hom}_{\mathrm{cnt}}(G_K, \{\pm 1\})$ correspond to a quadratic extension L/K.*

Then $\dim_{\mathbb{F}_2} Ш_{\mathrm{nd}}(A/K)[2] + \dim_{\mathbb{F}_2} Ш_{\mathrm{nd}}(A^\chi/K)[2] \equiv 0 \pmod 2$ *if and only if*

$$\sum_{v \text{ nonsplit in } L/K} \mathrm{inv}_v\, \mathfrak{g}(A/K_v, \lambda, \chi_v) = 0 \quad \text{in } \mathbb{Q}/\mathbb{Z}$$

where the local terms $\mathfrak{g}(A/K_v, \lambda, \chi_v) \in \mathrm{Br}(K_v)[2]$ are given in Definition 5.15. (Here χ_v denotes the restriction of χ to the completion K_v and $Ш_{\mathrm{nd}}(A/K)$ denotes the quotient of the Shafarevich–Tate group of A/K by its maximal divisible subgroup.)

In particular, Theorem 1.4 shows that the sum

$$\dim_{\mathbb{F}_2} Ш_{\mathrm{nd}}(A/K)[2] + \dim_{\mathbb{F}_2} Ш_{\mathrm{nd}}(A^\chi/K)[2] \pmod 2$$

is controlled by purely local behavior. In the case where A/K is the Jacobian of a curve, it is a result of Poonen and Stoll [1999, Corollary 12] that this is in fact true for the parity of $\dim_{\mathbb{F}_2} Ш_{\mathrm{nd}}(A/K)[2]$ itself, but whether or not this holds for an arbitrary principally polarized abelian variety remains open.

In general, the definition of the local terms $\mathfrak{g}(A/K_v, \lambda, \chi_v)$ appearing in Theorem 1.4 is somewhat involved but if the principal polarization λ on A/K_v is induced by a K_v-rational symmetric line bundle $\mathcal{L}_v$ then they take a simple form. Specifically, associated to $\mathcal{L}_v$ is a $\mathrm{Gal}(\overline{K_v}/K_v)$-invariant quadratic refinement q of the Weil pairing on $A[2]$ (we review this classical construction in Section 4B). As a consequence, $\mathrm{Gal}(\overline{K_v}/K_v)$ acts on $A[2]$ through the orthogonal group $O(q)$. In particular we obtain a

quadratic character ψ_v of K_v as the composition

$$\psi_v : \mathrm{Gal}(\overline{K_v}/K_v) \to O(q)/\mathrm{SO}(q) \cong \{\pm 1\}.$$

We then have

$$\mathfrak{g}(A/K_v, \lambda, \chi_v) = \chi_v \cup \psi_v \in \mathrm{Br}(K_v).$$

This allows the explicit evaluation of $\mathfrak{g}(A/K_v, \lambda, \chi_v)$ for archimedean places and for nonarchimedean places $v \nmid 2$ at which A has good reduction (Proposition 5.16). The implications for arithmetic of the difference in characteristic 2 between quadratic forms and symmetric bilinear pairings will be a recurring theme throughout this paper.

Theorem 1.4 may also be used to prove the analogue of Theorem 1.1 for the parity of the dimension of the 2-torsion of the Shafarevich–Tate group in the family of quadratic twists of a principally polarized abelian variety. This quantifies the failure of the Shafarevich–Tate group to have square order in the family of quadratic twists. See Theorem 10.27 for the precise statement.

To explain the remaining results of the paper we briefly indicate how we prove Theorem 1.1. As in [Klagsbrun et al. 2013], which proves the elliptic curve case, we deduce Theorem 1.1 from a more general theorem that determines the distribution of parities of ranks of certain Selmer groups $\mathrm{Sel}(T, \chi)$ associated to a finite dimensional $\mathbb{F}_p$-vector space T equipped with a $\mathrm{Gal}(\overline{K}/K)$-action, an alternating pairing, and abstract "twisting data". The general result is Theorem 7.4, the case $\dim T = 2$ of which combines Theorem 7.6 and Theorem 8.2 of [loc. cit.]. Taking $T = A[2]$ along with the Weil pairing and the twisting data detailed in Section 10 recovers Theorem 1.1.

On the other hand, taking $p > 2$ and $T = A[p]$ for a principally polarized abelian variety A/K, along with the twisting data described in Section 11, enables us to prove an analogue of Theorem 1.1 which applies to Selmer groups of certain p-cyclic twists of A^{p-1} (again, the case where A is an elliptic curve is shown by Klagsbrun, Mazur and Rubin [2013]). To state the result, let $\mathcal{C}(K)$ and $\mathcal{C}(K, X)$ for $p > 2$ be defined in the identical way to $p = 2$, replacing $\mathrm{Hom}_{\mathrm{cnt}}(\mathrm{Gal}(\overline{K}/K), \{\pm 1\})$ (the group of quadratic characters) with the group $\mathrm{Hom}_{\mathrm{cnt}}(\mathrm{Gal}(\overline{K}/K), \boldsymbol{\mu}_p)$ (of p-cyclic characters). For $\chi \in \mathcal{C}(K)$ nontrivial, let $L = \overline{K}^{\ker(\chi)}$ and denote by A^χ/K the $p-1$-dimensional abelian variety defined as the kernel of the norm homomorphism $\mathrm{Res}_{L/K} A \to A$ (here $\mathrm{Res}_{L/K}$ denotes the restriction of scalars from L to K). There is a natural inclusion of $\mathbb{Z}[\boldsymbol{\mu}_p]$ into $\mathrm{End}_K(A^\chi)$ and, in this way, any generator π of the unique prime of $\mathbb{Z}[\boldsymbol{\mu}_p]$ lying over p yields a self-isogeny of A^χ. Denote by $\mathrm{Sel}_\pi(A^\chi/K)$ the associated π-Selmer group which may be shown to be independent of the choice of π (see Section 11 for more details of the above constructions). We then have:

Theorem 1.5. *Let p be an odd prime, K a number field, A/K a principally polarized abelian variety, and Σ the set consisting of all archimedean places of K, all places of bad reduction for A, and all places dividing p. Define $\epsilon : \mathrm{Gal}(K(A[p])/K) \to \{\pm 1\}$ by $\sigma \mapsto (-1)^{\dim_{\mathbb{F}_p} A[p]^\sigma}$.*

(i) *Suppose ϵ is trivial when restricted to $\mathrm{Gal}(K(A[p])/K(\boldsymbol{\mu}_p))$. Then for all sufficiently large X,*

$$\frac{|\{\chi \in \mathcal{C}(K, X) : \dim_{\mathbb{F}_p} \mathrm{Sel}_\pi(A^\chi/K) \text{ is even}\}|}{|\mathcal{C}(K, X)|} = \frac{1 + (-1)^{\dim_{\mathbb{F}_p} \mathrm{Sel}_p(A/K)} \cdot \delta}{2}$$

where δ is an explicit finite product of local terms δ_v (see the statement of Corollary 11.6 for their definition). Moreover (unlike the case $p = 2$) δ is always nonzero.

(ii) *If ϵ is nontrivial when restricted to* $\mathrm{Gal}(K(A[p])/K(\boldsymbol{\mu}_p))$ *then*

$$\lim_{X\to\infty} \frac{|\{\chi \in \mathcal{C}(K, X) : \dim_{\mathbb{F}_p} \mathrm{Sel}_\pi(A^\chi/K) \text{ is even}\}|}{|\mathcal{C}(K, X)|} = \frac{1}{2}.$$

When A is an elliptic curve and $p > 2$, ϵ is nontrivial when restricted to $\mathrm{Gal}(K(A[p])/K(\boldsymbol{\mu}_p))$ if and only if p divides $[K(A[p]) : K]$ (see [Klagsbrun et al. 2013, Lemma 4.3]), so now both cases of Theorem 1.5 can occur. In particular, we see that allowing the dimension of A to be arbitrary uncovers a more uniform picture between $p = 2$ and $p > 2$ than was visible for elliptic curves. See Remark 11.8 for a discussion on conditions on the $\mathrm{Gal}(\bar{K}/K)$-action on $A[p]$ which result in case (i) and (ii), respectively, of Theorem 1.5. We simply note here that if the Galois action on $A[p]$ is as large as possible, so that $\mathrm{Gal}(K(A[p])/K)$ is isomorphic to the general symplectic group $\mathrm{GSp}_{2g}(\mathbb{F}_p)$ for $g = \dim A$, then case (ii) applies.

Finally, we remark that a key step in proving Theorem 7.4 (the version of Theorems 1.1 and 1.5 for general T) is, for a character χ, to describe the quantity

$$\dim_{\mathbb{F}_p} \mathrm{Sel}(T, \chi) - \dim_{\mathbb{F}_p} \mathrm{Sel}(T, \mathbb{1}) \pmod 2$$

as a sum of local terms (see Theorem 6.12). Upon taking $T = A[2]$ for a principally polarized abelian variety A/K one obtains (Theorem 10.12) a local formula for the difference between the parity of the 2-Selmer rank of A/K and the 2-Selmer rank of the quadratic twist A^χ/K. This generalizes a theorem of Kramer [1981, Theorem 1] for elliptic curves, and Yu [2016, Theorem 5.11] for Jacobians of odd degree hyperelliptic curves. Combining this with Theorem 1.4, one obtains (Theorem 10.20) a purely local formula for the parity of the 2^∞-Selmer rank of A over the quadratic extension cut out by χ. Such local formulae for (the parity of) 2^∞-Selmer ranks have applications to the 2-parity conjecture and we plan to examine this in future work.

Layout of the paper. In Section 2 we review some standard results in group cohomology which will be used in the sequel. In Section 3 we review and study quadratic forms on finite dimensional $\mathbb{F}_2$-vector spaces. The main result is Proposition 3.9 which forms a key technical step in the proof of Theorem 1.4. Section 4 recalls the constructions of certain quadratic forms associated to abelian varieties and examines how these behave under quadratic twist. Of particular importance is Lemma 4.20 which plays a crucial role in associating twisting data to the group of 2-torsion points of a principally polarized abelian variety. Theorem 1.4 is proven in Section 4. The analogue of Theorems 1.1 and 1.5 for general T is proven in Sections 6–9 which broadly follow the layout and strategy of [Klagsbrun et al. 2013, §3–4 and §6–8]. Specifically, in Section 6 we recall the notions of metabolic structure and twisting data from [loc. cit.] and generalize them to arbitrary (finite) dimensional $\mathbb{F}_p$-vector spaces, as well as defining the associated Selmer groups. Section 7 states the main result, Theorem 7.4, and proves the analogue of case (i) of Theorems 1.1 and 1.5 in this setting. Section 8 uses class field theory to produce certain global characters

with specified local behavior and is a more or less direct generalization of [Klagsbrun et al. 2013, §6], albeit with different proofs. The results of Section 8 are then applied in Section 9 to prove the remaining cases of Theorem 7.4. Section 10 associates a metabolic structure and twisting data to the 2-torsion in a principally polarized abelian variety and deduces Theorems 1.1 and 1.2. Finally, Section 11 associates a metabolic structure and twisting data to the p-torsion in a principally polarized abelian variety for p odd and deduces Theorem 1.5.

Notation. For a group G acting on an abelian group M, for $\sigma \in G$ we write

$$M^\sigma := \{m \in M : \sigma(m) = m\}.$$

For a field F we denote its separable closure by $\bar{F}$, its absolute Galois group by G_F and, for p different from the characteristic of F, we denote by $\boldsymbol{\mu}_p$ the G_F-module of p-th roots of unity in $\bar{F}$. We denote by $\mathrm{Br}(F)$ the Brauer group of F.

For an abelian variety A/F we write $A^\vee/F$ for the dual of A. A *principally polarized abelian variety* over F is a pair $(A/F, \lambda)$ consisting of an abelian variety A/F and a principal polarization $\lambda : A \to A^\vee$ defined over F. For a quadratic character $\chi \in \mathrm{Hom}_{\mathrm{cnt}}(G_F, \{\pm 1\})$ the *quadratic twist* of A by χ is the pair (A^χ, ψ) consisting of an abelian variety A^χ/F and an $\bar{F}$-isomorphism $\psi : A \to A^\chi$ such that $\psi^{-1}\psi^\sigma = [\chi(\sigma)]$ for all $\sigma \in G_F$.

For a number field K we denote by M_K the set of places of K and write K_v for the completion of K at $v \in M_K$. We denote by $\mathrm{inv}_v : \mathrm{Br}(K_v) \to \mathbb{Q}/\mathbb{Z}$ the local invariant map and, if v is nonarchimedean, denote by K_v^{ur} the maximal unramified extension of K_v. We implicitly fix embeddings $\bar{K} \hookrightarrow \overline{K_v}$ for each $v \in M_K$ and view G_{K_v} as a subgroup of G_K for each v. In particular, for a (finite) Galois extension L/K of number fields and a nonarchimedean place $v \in M_K$ unramified in L/K we have a well defined Frobenius element Frob_v in $\mathrm{Gal}(L/K)$.

For a G_K-module M, the injections $G_{K_v} \hookrightarrow G_K$ induce restriction maps on cohomology $H^i(K, M) \to H^i(K_v, M)$ for each $i \geq 0$ and $v \in M_K$. For a cocycle ξ we write ξ_v for its restriction to K_v (see Section 2 for our notation and conventions concerning group cohomology). We define, for v a nonarchimedean place of K,

$$H^i_{\mathrm{ur}}(K_v, M) := \ker(H^i(K_v, M) \xrightarrow{\mathrm{res}} H^i(K_v^{\mathrm{ur}}, M)).$$

2. Group cohomology and group extensions

In the following sections we will make several computations involving group cohomology. Here we set up the relevant notation and review some basic results. All material in this section is standard; see e.g., [Atiyah and Wall 1967].

2A. ***Group cohomology.*** Let G be a finite group and M a G-module. For $i \geq 0$ we write $C^i(G, M)$ for the group of i-cochains with values in M and $d : C^i(G, M) \to C^{i+1}(G, M)$ for the usual differential.

When $i = 0$ we have $(dm)(g) = gm - m$ for $m \in M = C^0(G, M)$ and $g \in G$, and when $i = 1$ we have

$$(df)(g, h) = f(g) + gf(h) - f(gh)$$

for $f \in C^1(G, M)$ and $g, h \in G$. We write $Z^i(G, M)$ and $B^i(G, M)$ for the group of i-cocycles and i-coboundaries, respectively, with values in M. We will always think of the i-th-cohomology group $H^i(G, M)$ as the quotient $Z^i(G, M)/B^i(G, M)$. When making computations involving group cohomology, we'll make the convention that fraktur letters such as $\mathfrak{a}$, $\mathfrak{b}$ etc. denote cohomology classes and that the corresponding lower case Roman letters a, b etc., denote cocycles representing these cohomology classes. More generally, if G is a profinite group we consider continuous cochains, cocycles and coboundaries, using the same notation and conventions to talk about them.

2B. ***Cup product on cochains.*** Let G be a finite (or profinite) group and let M and N be G-modules. Then for $i, j \geq 0$ the *cup-product* map

$$\cup : C^i(G, M) \times C^j(G, N) \to C^{i+j}(G, M \otimes N)$$

is defined by

$$(a \cup b)(g_1, \ldots, g_{i+j}) = a(g_1, \ldots, g_i) \otimes g_1 \cdots g_i b(g_{i+1}, \ldots, g_{i+j}).$$

For $a \in C^i(G, M)$ and $b \in C^j(G, N)$ we have the equality

$$d(a \cup b) = da \cup b + (-1)^i a \cup db \tag{2.1}$$

inside $C^{i+j+1}(G, M \otimes N)$.

For $i, j \geq 0$ the cup product map above induces a cup product map on cohomology

$$\cup : H^i(G, M) \times H^j(G, N) \to H^{i+j}(G, M \otimes N)$$

which satisfies $\mathfrak{a} \cup \mathfrak{b} = (-1)^{ij} \mathfrak{b} \cup \mathfrak{a}$.

2C. ***Group extensions.*** Let G be a finite group and M an abelian group with trivial G-action. In what follows we write the group law on G multiplicatively and the group law on M additively. Let $\mathfrak{a} \in H^2(G, M)$ and a be a 2-cocycle representing $\mathfrak{a}$. Define a group structure on the set $G \times M$ by the rule

$$(g, m) \cdot (g', m') = (gg', m + m' + a(g, g'))$$

and let E_a denote the resulting group. The maps $\alpha : M \to E_a$ and $\beta : E_a \to G$ defined by $m \mapsto (1, m - a(1, 1))$ and $(g, m) \mapsto g$ respectively give rise to the short exact sequence

$$0 \to M \xrightarrow{\alpha} E_a \xrightarrow{\beta} G \to 0$$

realizing E_a as a central extension of G by M. The isomorphism class of this extension is independent of the choice of cocycle representing $\mathfrak{a}$ and the sequence splits if and only if $\mathfrak{a}$ is the trivial class in $H^2(G, M)$. More specifically, let $s : G \to E_a$ denote the set section $g \mapsto (g, 0)$ to β. Then if

$\phi : E_a \to M$ is a homomorphism splitting the exact sequence (i.e., giving a section to α) then the function $f = \phi \circ s \in C^1(G, M)$ is a 1-cochain satisfying $df = a$.

Remark 2.2. The above correspondence in fact gives rise to a bijection between elements of $H^2(G, M)$ and the set of isomorphism classes of central extensions of G by M, and one can generalize this correspondence to include the case where the action of G on M is nontrivial (though now the relevant extensions are, in general, no longer central). See [Atiyah and Wall 1967, §2] for more details.

3. Quadratic forms on finite dimensional $\mathbb{F}_2$-vector spaces

The aim of this section is to prove Propositions 3.9 and 3.10 which are needed for the proof of Theorem 1.4. In Sections 3A, 3B and 3C we review the theory of quadratic forms on finite dimensional $\mathbb{F}_2$-vector spaces. The material in Sections 3A, and 3B is standard, see e.g., [Scharlau 1985, Section 9.4]. In Section 3C we review a construction due to Pollatsek [1971] (given in the discussion preceding Theorem 1.11 of that work) which we use in the proof of Proposition 3.9.

For the rest of this section fix a finite dimensional $\mathbb{F}_2$-vector space V equipped with a nondegenerate alternating pairing

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}_2$$

(so in particular $\dim V$ is even). We denote by $\mathrm{Sp}(V)$ the symplectic group of linear automorphisms of V preserving the pairing.

3A. *Quadratic refinements and the class* $\mathfrak{c} \in H^1(\mathrm{Sp}(V), V)$.

Definition 3.1 (quadratic refinement). A function $q : V \to \mathbb{F}_2$ is called a *quadratic refinement* of $\langle \cdot, \cdot \rangle$ if we have

$$q(v + v') + q(v) + q(v') = \langle v, v' \rangle$$

for all $v, v' \in V$.

Let $\mathcal{Q}$ denote the set of all quadratic refinements of $\langle \cdot, \cdot \rangle$. It is a principal homogeneous space for V where, for $v \in V$, we define $q + v \in \mathcal{Q}$ by setting

$$(q + v)(v') = q(v') + \langle v, v' \rangle$$

for $v' \in V$. The symplectic group $\mathrm{Sp}(V)$ acts on the set of quadratic refinements via $q \mapsto q \circ \sigma^{-1}$ (for $\sigma \in \mathrm{Sp}(V)$). This action is compatible with addition by elements of V and so associated to $\mathcal{Q}$ is a class

$$\mathfrak{c} \in H^1(\mathrm{Sp}(V), V).$$

Explicitly, picking a quadratic refinement q and defining $\lambda : V \to V^* := \mathrm{Hom}(V, \mathbb{F}_2)$ to be the map $v \mapsto \langle v, - \rangle$, the function $c_q : \mathrm{Sp}(V) \to V$ given by setting

$$c_q(\sigma) = \lambda^{-1}(q \circ \sigma^{-1} - q)$$

is a 1-cocycle representing $\mathfrak{c}$.

Remark 3.2. Let $\mathcal{A}lt$ denote the group of (possibly degenerate) alternating pairings on V under addition. It has an action of $\mathrm{Sp}(V)$ given by $\sigma \cdot \langle\langle \cdot, \cdot \rangle\rangle = \langle\langle \sigma^{-1}(\cdot), \sigma^{-1}(\cdot)\rangle\rangle$. Similarly, let $\mathcal{Q}uad$ denote the group of quadratic forms on V under addition which also carries an action of $\mathrm{Sp}(V)$ via $\sigma \cdot q = q \circ \sigma^{-1}$. Then we have a short exact sequence of $\mathrm{Sp}(V)$-modules

$$0 \to V^* \to \mathcal{Q}uad \to \mathcal{A}lt \to 0 \tag{3.3}$$

where the map $V^* \to \mathcal{Q}uad$ is inclusion and the map $\mathcal{Q}uad \to \mathcal{A}lt$ sends a quadratic form to its associated pairing. The associated long exact sequence for cohomology gives a map

$$\delta : H^0(\mathrm{Sp}(V), \mathcal{A}lt) \to H^1(\mathrm{Sp}(V), V^*).$$

Our pairing $\langle \cdot, \cdot \rangle$ is an element of $H^0(\mathrm{Sp}(V), \mathcal{A}lt)$ and the class $\mathfrak{c} \in H^1(\mathrm{Sp}(V), V)$ constructed above is the image of $\langle \cdot, \cdot \rangle$ under δ, once we use the map λ above to identify $H^1(\mathrm{Sp}(V), V)$ with $H^1(\mathrm{Sp}(V), V^*)$.

Remark 3.4. It is shown in [Pollatsek 1971, Theorems 4.1 and 4.4] that if $\dim(V) \geq 4$ then $H^1(\mathrm{Sp}(V),V) \cong \mathbb{Z}/2\mathbb{Z}$, generated by $\mathfrak{c}$.

3B. ***Orthogonal groups, special orthogonal groups and the Dickson homomorphism.*** For a given quadratic refinement q, denote by $O(q)$ the corresponding orthogonal group of linear automorphisms preserving q rather than just the pairing. The orthogonal group $O(q)$ has an index 2 subgroup $\mathrm{SO}(q)$ which is by definition the kernel of the Dickson homomorphism, whose definition we now recall. Let $C(q)$ denote the Clifford algebra associated to q (see [Scharlau 1985, Definition 9.2.1]), $C^0(q)$ its even graded subalgebra and $Z(q)$ the center of $C^0(q)$. Then $Z(q)$ is a rank 2 étale algebra over $\mathbb{F}_2$ (see Theorem 9.4.8 of [loc. cit.]). Since $O(q)$ acts naturally on $C(q)$ and preserves the grading, it acts on $Z(q)$ by $\mathbb{F}_2$-algebra homomorphisms. Noting that the automorphism group of any rank 2 étale algebra over $\mathbb{F}_2$ (or indeed any field) is canonically isomorphic to $\mathbb{Z}/2\mathbb{Z}$, we obtain a homomorphism $d_q : O(q) \to \mathbb{Z}/2\mathbb{Z}$, the *Dickson homomorphism*.

We will also need the following alternative characterization of the Dickson homomorphism.

Proposition 3.5. *Let q be a quadratic refinement of $\langle \cdot, \cdot \rangle$ and $\sigma \in O(q)$. Then*

$$d_q(\sigma) = \dim V^\sigma \pmod 2.$$

Proof. This is [Dye 1977, Theorem 3]. □

3C. ***An extension of the Dickson homomorphism to the full symplectic group.*** The following is a version of a construction due to Pollatsek [1971] which gives an extension of the Dickson homomorphism to the whole of $\mathrm{Sp}(V)$. We caution however that the resulting function $\mathrm{Sp}(V) \to \mathbb{Z}/2\mathbb{Z}$ is not a homomorphism (we cannot ask for this since for $\dim V \geq 6$ the group $\mathrm{Sp}(V)$ is simple).

Construction 3.6 (Pollatsek). Fix a quadratic refinement q of $\langle \cdot, \cdot \rangle$. Set $U = \mathbb{F}_2^2$ equipped with its unique nondegenerate alternating form $\langle \cdot, \cdot \rangle_U$. Further, let q_U denote the unique quadratic refinement of $\langle \cdot, \cdot \rangle_U$

with Arf invariant 1. Thus for $(\lambda, \lambda') \in U$ we have

$$q_U((\lambda, \lambda')) = \lambda + \lambda' + \lambda\lambda'.$$

Let $x = (1, 0)$ and $y = (0, 1)$ so that $q_U(x) = 1 = q_U(y)$ and $\langle x, y\rangle_U = 1$. Now let $W := V \oplus U$ be the orthogonal direct sum of V and U, so that W comes equipped with the quadratic form $q_W := q + q_U$, whose associated (nondegenerate, alternating) pairing is $\langle \cdot, \cdot \rangle_W := \langle \cdot, \cdot \rangle + \langle \cdot, \cdot \rangle_U$.

Now given $g = (\sigma, \alpha) \in \mathrm{Sp}(V) \times \mathbb{F}_2$, define the linear automorphism $\phi_q(g)$ of W by setting

$$\phi_q(g)(x) = x \quad \text{and} \quad \phi_q(g)(y) = \alpha x + c_q(\sigma) + y,$$

and for $v \in V$,

$$\phi_q(g)(v) = \sigma(v) + \langle c_q(\sigma), \sigma(v)\rangle x$$

and extending linearly.

A key property of this construction, as shown in the discussion preceding [Pollatsek 1971, Theorem 1.11], is that for each $g \in \mathrm{Sp}(V) \times \mathbb{F}_2$ we have $\phi_q(g) \in O(q_W)$. Moreover, Pollatsek shows in [loc. cit.] that for each $\sigma \in \mathrm{Sp}(V)$, there is a unique $\alpha(\sigma) \in \mathbb{F}_2$ such that $\phi_q((\sigma, \alpha(\sigma))) \in \mathrm{SO}(q_W)$. One has $\alpha(\sigma) = d_q(\sigma)$ for all $\sigma \in O(q)$, so the map $\sigma \mapsto \alpha(\sigma)$ gives an extension of the Dickson homomorphism to the full symplectic group $\mathrm{Sp}(V)$.

3D. ***Triviality of*** $\mathfrak{c} \cup \mathfrak{c}$**.** The pairing $\langle \cdot, \cdot \rangle$ induces a cup-product map

$$\cup : H^1(\mathrm{Sp}(V), V) \times H^1(\mathrm{Sp}(V), V) \to H^2(\mathrm{Sp}(V), \mathbb{F}_2).$$

We now use the construction of the previous subsection to analyze the element $\mathfrak{c} \cup \mathfrak{c} \in H^2(\mathrm{Sp}(V), \mathbb{F}_2)$.

Notation 3.7. Given a quadratic refinement $q \in \mathcal{Q}$, let E_q denote the central extension of $\mathrm{Sp}(V)$ by $\mathbb{F}_2$ corresponding to the 2-cocycle $c_q \cup c_q$, so that as a set $E_q = \mathrm{Sp}(V) \times \mathbb{F}_2$, and is equipped with the group structure

$$(\sigma, \alpha) \cdot (\sigma', \alpha') = (\sigma\sigma', \alpha + \alpha' + (c_q \cup c_q)(\sigma, \sigma')).$$

We then have:

Lemma 3.8. *The function ϕ_q of Construction 3.6 is a homomorphism $E_q \to O(q_W)$.*

Proof. As above, ϕ_q gives a map from E_q into $O(q_W)$. An easy computation shows additionally that it is a homomorphism. □

We may now prove the main result of the section.

Proposition 3.9. *For each quadratic refinement $q \in \mathcal{Q}$ there is a unique function $f_q : \mathrm{Sp}(V) \to \mathbb{F}_2$ such that $df_q = c_q \cup c_q \in Z^2(\mathrm{Sp}(V), \mathbb{F}_2)$ and such that the restriction of f_q to the orthogonal group $O(q)$ is the Dickson homomorphism. In particular, we have*

$$\mathfrak{c} \cup \mathfrak{c} = 0 \in H^2(\mathrm{Sp}(V), \mathbb{F}_2).$$

Proof. We first show uniqueness. If f_q' is another function with $df_q' = c_q \cup c_q$ then the difference $f_q - f_q'$ is a homomorphism from $\mathrm{Sp}(V)$ to $\mathbb{F}_2$. If $\dim V \geq 6$ then $\mathrm{Sp}(V)$ is simple and hence $f_q = f_q'$. If $\dim V$ (which is necessarily even) is 2 or 4 then $\mathrm{Sp}(V)$ has a unique index 2 subgroup and hence a unique nontrivial homomorphism to $\mathbb{F}_2$. In each case this homomorphism is nontrivial when restricted to $O(q)$ for each quadratic refinement q, whence the result.

In the notation of Construction 3.6, associated to q_W is the Dickson homomorphism

$$d_{q_W} : O(q_W) \to \mathbb{F}_2.$$

We claim that $d_{q_W} \circ \phi_q : E_q \to \mathbb{F}_2$ gives a section to the map $\mathbb{F}_2 \to E_q$ sending α to $(1, \alpha)$, thus splitting the extension E_q. Indeed, let $\alpha \in \mathbb{F}_2$. Then $\phi_q((1, \alpha)) = \mathrm{id}_V \oplus m_\alpha$ where $m_\alpha \in O(q_U)$ is defined by $m_\alpha(x) = x$, $m_\alpha(y) = \alpha x + y$. One sees (either using the definition in terms of Clifford algebras, or by applying Proposition 3.5) that $\mathrm{id}_V \oplus m_\alpha$ is in $\mathrm{SO}(q_W)$ if and only if $\alpha = 0$, whence $d_{q_W}((1, \alpha)) = \alpha$ as desired.

It now follows that the function $f_q : \mathrm{Sp}(V) \to \mathbb{F}_2$ defined by $f_q(\sigma) = (d_{q_W} \circ \phi_q)((\sigma, 0))$ satisfies $df_q = c_q \cup c_q$ (see Section 2C and note that $(c_q \cup c_q)(1, 1) = 0$).

It remains to show that the restriction of f_q to $O(q)$ is the Dickson homomorphism d_q. To see this note that for any $\sigma \in O(q)$ we have $c_q(\sigma) = 0$ and so

$$\phi_q((\sigma, 0)) = \sigma \oplus \mathrm{id}_U .$$

Since this is in $\mathrm{SO}(q_W)$ if and only if σ is in $\mathrm{SO}(q)$ (again by looking at Clifford algebras or using Proposition 3.5), we have the claim. □

We now describe how f_q changes upon changing the quadratic refinement q.

Proposition 3.10. *Let q and q' be two quadratic refinements of $\langle \cdot, \cdot \rangle$ and let $v \in V$ be such that $q' = q + v$, so that $c_{q'} = c_q + dv$. Then we have*

$$f_{q'} = f_q + c_q \cup v + v \cup c_q + v \cup dv$$

as cochains in $C^1(\mathrm{Sp}(V), \mathbb{F}_2)$.

Proof. One readily computes

$$d(f_q + c_q \cup v + v \cup c_q + v \cup dv) = c_{q'} \cup c_{q'},$$

so it remains to show that the restriction of $f_q + c_q \cup v + v \cup c_q + v \cup dv$ to $O(q')$ is the Dickson homomorphism $d_{q'}$. To do this we'll use the characterization of the Dickson homomorphism given in Proposition 3.5.

Fix $\sigma \in O(q')$. Then $c_q(\sigma) = (dv)(\sigma)$. In the notation of Construction 3.6, given $w \in W$ and writing $w = z + \epsilon_1 x + \epsilon_2 y$ with $z \in V$ and $\epsilon_1, \epsilon_2 \in \mathbb{F}_2$, one sees that w is fixed by $\phi_q((\sigma, 0))$ if and only if

$$\sigma(z) - z = \epsilon_2 (dv)(\sigma) \tag{3.11}$$

and

$$\langle (dv)(\sigma), \sigma(z)\rangle = 0. \tag{3.12}$$

Now (3.11) is equivalent to $z = z' + \epsilon_2 v$ for some $z' \in V^\sigma$. If z has this form, then using invariance of z' under σ one computes

$$\langle (dv)(\sigma), \sigma(z)\rangle = \epsilon_2 \langle \sigma(v), v\rangle.$$

Thus if $\langle \sigma(v), v\rangle = 0$ then the second condition (3.12) is redundant, whilst if $\langle \sigma(v), v\rangle = 1$ then it may be replaced with the condition $\epsilon_2 = 0$. We conclude that

$$\dim W^{\phi_q((\sigma,0))} \equiv \dim V^\sigma + \langle \sigma(v), v\rangle \pmod 2$$

and hence (using Proposition 3.5)

$$f_q(\sigma) = d_{q'}(\sigma) + \langle \sigma(v), v\rangle = d_{q'}(\sigma) + (v \cup dv)(\sigma).$$

Thus the restriction of f_q to $O(q')$ is equal to $d_{q'} + v \cup dv$. Noting also that the restriction of c_q to $O(q')$ is equal to dv the result follows easily. □

Remark 3.13. Let $\tilde{V}$ denote the group whose underlying set is $V \times \mathbb{F}_2$, endowed with the group law

$$(v, \alpha) \cdot (v', \alpha') = (v + v', \alpha + \alpha' + \langle v, v'\rangle).$$

Then $\tilde{V}$ sits in a short exact sequence

$$0 \to \mathbb{F}_2 \to \tilde{V} \to V \to 0, \tag{3.14}$$

the map $\mathbb{F}_2 \to \tilde{V}$ sending α to $(0, \alpha)$ and the map $\tilde{V} \to V$ being projection onto the first factor. Making $\mathrm{Sp}(V)$ act trivially on $\mathbb{F}_2$ and diagonally on $\tilde{V}$ this sequence becomes an exact sequence of $\mathrm{Sp}(V)$-modules. Using the relation $df_q = c_q \cup c_q$ one can show that for each quadratic refinement q the function $\tilde{c}_q : \mathrm{Sp}(V) \to \tilde{V}$ defined by

$$\tilde{c}_q(\sigma) = (c_q(\sigma), f_q(\sigma))$$

is a 1-cocycle. One may then use the relationship between f_q and f'_q given in Proposition 3.10 to show that the class $\tilde{\mathfrak{c}}$ of $\tilde{c}_q$ in $H^1(\mathrm{Sp}(V), \tilde{V})$ does not depend on q so that the results of this section prove that $\mathfrak{c} \in H^1(\mathrm{Sp}(V), V)$ admits a canonical lift to $H^1(\mathrm{Sp}(V), \tilde{V})$. (It is shown in [Poonen and Rains 2011, Corollary 2.8(b)] that the connecting homomorphism $H^1(\mathrm{Sp}(V), V) \to H^2(\mathrm{Sp}(V), \mathbb{F}_2)$ arising from (3.14) sends $\mathfrak{a} \in H^1(\mathrm{Sp}(V), V)$ to $\mathfrak{a} \cup \mathfrak{a}$, so that the triviality of $\mathfrak{c} \cup \mathfrak{c}$ is equivalent to the existence of some lift of $\mathfrak{c}$ to $H^1(\mathrm{Sp}(V), \tilde{V})$.)

4. Quadratic forms associated to abelian varieties

In this section we study the behavior under quadratic twist of certain quadratic forms associated to abelian varieties. Though several results in this section will be used in what follows, the most important is Lemma 4.20 which provides the technical input required to generalize [Yu 2016, Theorem 5.10] to the case

of arbitrary principally polarized abelian varieties (this is done in Lemma 10.6). Sections 4A–4C review some standard results in the theory of abelian varieties as can be found, for example, in [Mumford 1966].

For the rest of this section, fix a field F of characteristic 0 (which for applications will be either a number field or the completion of one). Let A/F be an abelian variety. For $x \in A(\bar{F})$ denote by τ_x the translation-by-x map $\tau_x : A \to A$.

4A. ***Line bundles and self-dual homomorphisms.*** Let $\mathcal{L}$ be a line bundle on $A/\bar{F}$. We denote by $\phi_{\mathcal{L}}$ the homomorphism $A \to A^\vee$ sending $x \in A(\bar{F})$ to the element of $A^\vee(\bar{F})$ corresponding to the line bundle $\tau_x^*\mathcal{L} \otimes \mathcal{L}^{-1}$. We write $K(\mathcal{L})$ for the kernel of $\phi_{\mathcal{L}}$. If $\mathcal{L}$ is ample then $K(\mathcal{L})$ is a finite subgroup of A.

We have a short exact sequence of G_F-modules

$$0 \to A^\vee(\bar{F}) \to \operatorname{Pic} A_{\bar{F}} \to \operatorname{Hom}_{\text{self-dual}}(A_{\bar{F}}, A^\vee_{\bar{F}}) \to 0, \tag{4.1}$$

the map $A^\vee(\bar{F}) \to \operatorname{Pic} A_{\bar{F}}$ being the natural inclusion and the map $\operatorname{Pic} A_{\bar{F}} \to \operatorname{Hom}_{\text{self-dual}}(A_{\bar{F}}, A^\vee_{\bar{F}})$ sending a line bundle $\mathcal{L}$ to $\phi_{\mathcal{L}}$. As in [Poonen and Rains 2011, §3.2], (4.1) induces a short exact sequence of G_F-modules

$$0 \to A^\vee[2] \to \operatorname{Pic}^{\text{sym}} A_{\bar{F}} \to \operatorname{Hom}_{\text{self-dual}}(A_{\bar{F}}, A^\vee_{\bar{F}}) \to 0, \tag{4.2}$$

where here $\operatorname{Pic}^{\text{sym}} A_{\bar{F}}$ denotes the group of symmetric line bundles on A (i.e., those satisfying $[-1]^*\mathcal{L} \cong \mathcal{L}$).

4B. ***Quadratic refinements of the Weil pairing on A[2].*** Let $(\cdot,\cdot)_{e_2} : A[2] \times A^\vee[2] \to \boldsymbol{\mu}_2$ denote the Weil pairing. It is bilinear, nondegenerate and G_F-equivariant. If $\lambda : A \to A^\vee$ is a self-dual homomorphism then it induces an alternating pairing

$$(\cdot,\cdot)_\lambda : A[2] \times A[2] \to \boldsymbol{\mu}_2$$

defined by $(a, b)_\lambda = (a, \lambda(b))_{e_2}$ for $a, b \in A[2]$. If λ is defined over F then $(\cdot,\cdot)_\lambda$ is G_F-invariant. In general, for a line bundle $\mathcal{L}$ on A set $(\cdot,\cdot)_{\mathcal{L}} := (\cdot,\cdot)_{\phi_{\mathcal{L}}}$.

Definition 4.3. Let $\mathcal{L}$ be a symmetric line bundle on A. Define the map $q_{\mathcal{L}} : A[2] \to \boldsymbol{\mu}_2$ as follows. Given $x \in A[2]$, we have $x^*[-1]^*\mathcal{L} = x^*\mathcal{L}$. In particular, the restriction of the normalized[1] isomorphism $\tau : \mathcal{L} \xrightarrow{\sim} [-1]^*\mathcal{L}$ to x is multiplication by an element $\eta_x \in \bar{F}^\times$ on $x^*\mathcal{L}$. One in fact has $\eta_x \in \boldsymbol{\mu}_2$ and we set $q_{\mathcal{L}}(x) := \eta_x$.

Remark 4.4. The map $q_{\mathcal{L}}$ defined above is denoted $e^{\mathcal{L}}_*$ in [Mumford 1966, fourth definition in §2].

The following well known lemma summarizes the properties of $q_{\mathcal{L}}$.

Lemma 4.5. *Let $\mathcal{L}$ be a symmetric line bundle on A. Then we have*:

(i) *If $\mathcal{L} \cong \mathcal{L}'$ then $q_{\mathcal{L}} = q_{\mathcal{L}'}$.*

[1] Writing $e \in A(\bar{F})$ for the identity section, an isomorphism $\tau : \mathcal{L} \xrightarrow{\sim} [-1]^*\mathcal{L}$ is called *normalized* if

$$e^*(\tau) : e^*\mathcal{L} \xrightarrow{\sim} e^*[-1]^*\mathcal{L} = e^*\mathcal{L}$$

is the identity. There is a unique such τ for each symmetric line bundle (see [Mumford 1966, §2]).

(ii) *The function $q_{\mathcal{L}}$ is a quadratic form on $A[2]$ (valued in $\boldsymbol{\mu}_2$) whose associated bilinear pairing is $(\cdot,\cdot)_{\mathcal{L}}$.*

(iii) *If $\mathcal{M}$ is another symmetric line bundle then $q_{\mathcal{L}\otimes\mathcal{M}} = q_{\mathcal{L}} \cdot q_{\mathcal{M}}$.*

Proof. Part (i) is immediate. For parts (ii) and (iii) see e.g., [Mumford 1966, §2; Poonen and Rains 2011, Proposition 3.2]. □

For a principal polarization $\lambda : A \to A^\vee$ defined over F, we can use Lemma 4.5 to give a geometric interpretation of the principal homogeneous space for $A[2]$ associated to the set of quadratic refinements of the Weil pairing $(\cdot,\cdot)_\lambda$ on $A[2]$.

Definition 4.6. Let $\lambda : A \to A^\vee$ be a self-dual homomorphism defined over F. We define $\mathfrak{c}_\lambda \in H^1(F, A^\vee[2])$ to be the image of λ under the connecting homomorphism in the long exact for Galois cohomology associated to (4.2). If λ is a principal polarization we will also, by an abuse of notation, write $\mathfrak{c}_\lambda$ for the element $\lambda^{-1}(\mathfrak{c}_\lambda) \in H^1(F, A[2])$.

Lemma 4.7. *Let $\lambda : A \to A^\vee$ be a principal polarization defined over F, so that $(\cdot,\cdot)_\lambda$ is a nondegenerate, G_F-equivariant, alternating pairing on $A[2]$. Then G_F acts on $A[2]$ through the symplectic group $\mathrm{Sp}(A[2])$ associated to the pairing $(\cdot,\cdot)_\lambda$. Let $\mathfrak{c} \in H^1(F, A[2])$ be the cohomology class associated to the set of quadratic refinements of $(\cdot,\cdot)_\lambda$ as in Section 3A.*

Then we have the equality $\mathfrak{c} = \mathfrak{c}_\lambda$ inside $H^1(F, A[2])$.

Proof. We remark that this is implicit in [Poonen and Rains 2011, §3]. First note that by Lemma 4.5(ii), for any symmetric line bundle $\mathcal{L}$ for which $\lambda = \phi_{\mathcal{L}}$, the function $q_{\mathcal{L}}$ is a quadratic refinement of $(\cdot,\cdot)_\lambda$. The result now follows either by an explicit computation using the association $\mathcal{L} \mapsto q_{\mathcal{L}}$ or, more conceptually, from the long exact sequences for cohomology associated to the commutative diagram (16) of [Poonen and Rains 2011, §3.4], the top row of which is our sequence (4.2) and the bottom row of which is the exact sequence (3.3) of Remark 3.2. □

4C. ***Theta groups.*** In this subsection we suppose that $\mathcal{L}$ is an ample line bundle on A so that $K(\mathcal{L})$ is finite. We recall the definition of the Theta group associated to $\mathcal{L}$ (see [Mumford 1966] for more details of what follows).

Definition 4.8. The *Theta group* $\mathcal{G}(\mathcal{L})$ associated to $\mathcal{L}$ is the set of pairs (x, φ) where $x \in K(\mathcal{L})$ and φ is an isomorphism $\varphi : \mathcal{L} \xrightarrow{\sim} \tau_x^*\mathcal{L}$ (over $\bar{F}$). The group operation is given by

$$(x, \varphi) \cdot (x', \varphi') = (x + x', \tau_{x'}^*(\varphi) \circ \varphi').$$

Remark 4.9. If $\mathcal{L} \cong \mathcal{L}'$ then fixing an isomorphism $\alpha : \mathcal{L} \xrightarrow{\sim} \mathcal{L}'$ we obtain an isomorphism $\mathcal{G}(\mathcal{L}) \xrightarrow{\sim} \mathcal{G}(\mathcal{L}')$ given by

$$(x, \varphi) \mapsto (x, \tau_x^*(\alpha) \circ \varphi \circ \alpha^{-1})$$

which is independent of α (since any two choices differ by a scalar). As such, $\mathcal{G}(\mathcal{L})$ is canonically isomorphic to $\mathcal{G}(\mathcal{L}')$.

Remark 4.10. The group $\mathcal{G}(\mathcal{L})$ sits in a short exact sequence

$$0 \to \bar{F}^\times \to \mathcal{G}(\mathcal{L}) \to K(\mathcal{L}) \to 0, \tag{4.11}$$

the map $\mathcal{G}(\mathcal{L}) \to K(\mathcal{L})$ being projection onto the first factor and the map $\bar{F}^\times \to \mathcal{G}(\mathcal{L})$ sending $\eta \in \bar{F}^\times$ to the pair (0, multiplication by η).

Lemma 4.12. *We have the following functorial properties of* $\mathcal{G}$:

(i) *Let A/F and B/F be abelian varieties, let $\mathcal{L}$ be an ample line bundle on B and let $f : A \to B$ be an isomorphism. Then the map $\tilde{f} : \mathcal{G}(f^*\mathcal{L}) \xrightarrow{\sim} \mathcal{G}(\mathcal{L})$ given by*

$$(x, \varphi) \mapsto (f(x), (f^{-1})^*(\varphi))$$

is an isomorphism making the diagram

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \bar{F}^\times & \longrightarrow & \mathcal{G}(f^*\mathcal{L}) & \longrightarrow & K(f^*\mathcal{L}) & \longrightarrow & 0 \\ & & \| & & \downarrow{\scriptstyle \tilde{f}} & & \downarrow{\scriptstyle f} & & \\ 0 & \longrightarrow & \bar{F}^\times & \longrightarrow & \mathcal{G}(\mathcal{L}) & \longrightarrow & K(\mathcal{L}) & \longrightarrow & 0 \end{array}$$

commute.

(ii) *Gven abelian varieties A/F, B/F and C/F, isomorphisms $f_1 : A \to B$ and $f_2 : B \to C$ and an ample line bundle $\mathcal{L}$ on C, we have*

$$\widetilde{f_2 \circ f_1} = \tilde{f}_2 \circ \tilde{f}_1 : \mathcal{G}(f_1^* f_2^* \mathcal{L}) \xrightarrow{\sim} \mathcal{G}(\mathcal{L}).$$

Proof. In both cases this is a simple computation. We remark that we crucially require that f is an isomorphism in (i), the situation for a general homomorphism being more subtle. See, for example, [Mumford 1966, Proposition 2] and the surrounding discussion. □

4D. ***Theta groups in the main case of interest.*** Suppose that A is equipped with a fixed principal polarization $\lambda : A \to A^\vee$ defined over F and take $\mathcal{L} = (1, \lambda)^*\mathcal{P}$ where $\mathcal{P}$ is the Poincaré line bundle on $A \times A^\vee$ (here, for a homomorphism $\mu : A \to A^\vee$ we denote by $(1, \mu) : A \to A \times A^\vee$ the composition of the diagonal morphism $\Delta : A \to A \times A$ with the morphism $1 \times \mu : A \times A \to A \times A^\vee$). Then $\mathcal{L}$ is an F-rational, ample, symmetric line bundle on A such that $\phi_{\mathcal{L}} = 2\lambda$ (see [Poonen and Rains 2012, Remark 4.5]). In particular, we have $K(\mathcal{L}) = \ker(2\lambda) = A[2]$.

Since $[-1]^*\mathcal{L} \cong \mathcal{L}$ we have an induced automorphism $\widetilde{[-1]}$ of $\mathcal{G}(\mathcal{L})$ as in Lemma 4.12.

Lemma 4.13. *With $\mathcal{L} = (1, \lambda)^*\mathcal{P}$ as above, the automorphism $\widetilde{[-1]}$ of $\mathcal{G}(\mathcal{L})$ is trivial.*

Proof. By [Mumford 1966, Proposition 3], if $\mathcal{F}$ is any ample symmetric line bundle on A and $(x, \varphi) \in \mathcal{G}(\mathcal{F})$ is such that $x \in A[2]$, then the automorphism $\widetilde{[-1]}$ of $\mathcal{G}(\mathcal{F})$ sends (x, φ) to $(x, q_{\mathcal{F}}(x)\varphi)$.

In particular, since $K(\mathcal{L}) = A[2]$ in our case, it suffices to show that $q_{\mathcal{L}}$ is trivial. Pick a symmetric line bundle $\mathcal{M}$ such that $\lambda = \phi_{\mathcal{M}}$ (whilst it may not be possible to choose an F-rational such $\mathcal{M}$, this is always possible over $\bar{F}$). By standard properties of the Poincaré line bundle we have $(1 \times \phi_{\mathcal{M}})^*\mathcal{P} \cong$

$m^*\mathcal{M} \otimes p_1^*\mathcal{M}^{-1} \otimes p_2^*\mathcal{M}^{-1}$, where $m : A \times A \to A$ is addition and p_1 and p_2 denote projection onto the first and second factors. Pulling back along the diagonal morphism $\Delta : A \to A \times A$ we obtain

$$\mathcal{L} \cong [2]^*\mathcal{M} \otimes \mathcal{M}^{-2} \cong \mathcal{M}^2$$

where for the second isomorphism above we use the symmetry of $\mathcal{M}$ along with the fact that for any line bundle $\mathcal{F}$ on A we have $[2]^*\mathcal{F} \cong \mathcal{F}^3 \otimes [-1]^*\mathcal{F}$ (see e.g., [Milne 1986, Corollary 6.6]). By Lemma 4.5(iii) we conclude that $q_{\mathcal{L}} = (q_{\mathcal{M}})^2 = 1$ as desired. □

Remark 4.14. As $\mathcal{L}$ is F-rational, the group $\mathcal{G}(\mathcal{L})$ carries a natural G_F-action. Explicitly, for $\sigma \in G_F$ and $(x, \varphi) \in \mathcal{G}(\mathcal{L})$, we have

$$\sigma \cdot (x, \varphi) = (\sigma(x), \sigma^*(\varphi)) \in \mathcal{G}(\sigma^*\mathcal{L}) = \mathcal{G}(\mathcal{L})$$

where for the equality $\mathcal{G}(\sigma^*\mathcal{L}) = \mathcal{G}(\mathcal{L})$ we combine Remark 4.9 with the assumption that $\mathcal{L}$ is F-rational. In particular, the exact sequence of Remark 4.10 becomes a short exact sequence of G_F groups

$$0 \to \bar{F}^\times \to \mathcal{G}(\mathcal{L}) \to A[2] \to 0 \tag{4.15}$$

(we caution here that $\mathcal{G}(\mathcal{L})$ is nonabelian). This short exact sequence will be important in what follows. More specifically, as in [Poonen and Rains 2012, Corollary 4.7], the associated connecting map $H^1(F, A[2]) \to H^2(F, \bar{F}^\times)$ is a quadratic form whose associated bilinear pairing is that arising from cup-product and the Weil pairing $(\cdot, \cdot)_\lambda : A[2] \times A[2] \to \boldsymbol{\mu}_2 \hookrightarrow \bar{F}^\times$.

4E. ***Quadratic twists.*** Maintaining the notation of Section 4D (so in particular $\mathcal{L} = (1, \lambda)^*\mathcal{P}$) let $\chi : G_F \to \boldsymbol{\mu}_2$ be a quadratic character. Write (A^χ, ψ) for the quadratic twist of A by χ (so that $\psi : A \to A^\chi$ is an $\bar{F}$-isomorphism with $\psi^{-1} \circ \psi^\sigma = [\chi(\sigma)]$ for all $\sigma \in G_F$). We now consider the effect of quadratic twisting on the constructions appearing earlier in this section. Note that ψ restricts to a G_F-equivariant isomorphism $A[2] \xrightarrow{\sim} A^\chi[2]$.

Lemma 4.16. *The morphism $\lambda_\chi := (\psi^\vee)^{-1}\lambda\psi^{-1} : A^\chi \to A^{\chi\vee}$ is a principal polarization defined over F.*

Proof. This is a manifestation of the fact that $[-1]^*$ acts trivially on the Néron–Severi group. More precisely, one computes immediately that λ_χ is defined over F, and it's a polarization since if $\mathcal{M}$ is a line bundle on A (not necessarily F-rational) such that $\lambda = \phi_{\mathcal{M}}$ then one has $\lambda_\chi = \phi_{\mathcal{M}_\chi}$ where $\mathcal{M}_\chi = (\psi^{-1})^*\mathcal{M}$. □

More generally, we have:

Lemma 4.17. *We have a commutative diagram of G_F-modules*

$$\begin{array}{ccccccccc}
0 & \longrightarrow & A^{\chi\vee}[2] & \longrightarrow & \mathrm{Pic}^{\mathrm{sym}} A^\chi_{\bar{F}} & \longrightarrow & \mathrm{Hom}_{\textit{self-dual}}(A^\chi_{\bar{F}}, A^{\chi\vee}_{\bar{F}}) & \longrightarrow & 0 \\
 & & \downarrow{\scriptstyle \psi^\vee} & & \downarrow{\scriptstyle \psi^*} & & \downarrow & & \\
0 & \longrightarrow & A^\vee[2] & \longrightarrow & \mathrm{Pic}^{\mathrm{sym}} A_{\bar{F}} & \longrightarrow & \mathrm{Hom}_{\textit{self-dual}}(A_{\bar{F}}, A^\vee_{\bar{F}}) & \longrightarrow & 0,
\end{array}$$

where the rightmost vertical map sends μ to $\psi\mu\psi^\vee$.

Proof. As with Lemma 4.16 this follows from an explicit computation, and results from the fact that $[-1]$ acts trivially on each group appearing. □

Corollary 4.18. *Let $\overline{\psi}^{-1}$ denote the isomorphism $H^1(F, A^\chi[2]) \to H^1(F, A[2])$ induced by ψ^{-1} and let $\mathfrak{c}_\lambda \in H^1(F, A[2])$ and $\mathfrak{c}_{\lambda_\chi} \in H^1(F, A^\chi[2])$ be the cohomology class associated to λ and λ_χ, respectively, as in Definition 4.6. Then we have $\overline{\psi}^{-1}(\mathfrak{c}_{\lambda_\chi}) = \mathfrak{c}_\lambda$.*

Proof. This follows immediately from the long exact sequences for cohomology associated to the commutative diagram of Lemma 4.17. □

We now consider the effect of quadratic twisting on the Theta group associated to $\mathcal{L} = (1, \lambda)^*\mathcal{P}$.

Lemma 4.19. *Let $\mathcal{L} = (1, \lambda)^*\mathcal{P}$, write $\mathcal{P}_\chi$ for the Poincaré line bundle on $A^\chi \times A^{\chi\vee}$ and define $\mathcal{L}_\chi := (1, \lambda_\chi)^*\mathcal{P}_\chi$. Then $\psi^*\mathcal{L}_\chi \cong \mathcal{L}$.*

Proof. Standard properties of the Poincaré line bundle (see e.g., [Milne 1986, §11]) give

$$(1 \times \psi^\vee)^*\mathcal{P} \cong (\psi \times 1)^*\mathcal{P}_\chi$$

as line bundles on $A \times A^{\chi\vee}$. Since $\psi^\vee$ is an isomorphism we obtain

$$\mathcal{L} = (1, \lambda)^*\mathcal{P} \cong (1, \lambda)^*(1 \times (\psi^\vee)^{-1})^*(\psi \times 1)^*\mathcal{P}_\chi.$$

The right-hand side of the above expression is easily seen to be equal to

$$\psi^*\Delta^*(1 \times \lambda_\chi)^*\mathcal{P}_\chi = \psi^*\mathcal{L}_\chi$$

as desired (here $\Delta : A \to A \times A$ is the diagonal morphism). □

Lemma 4.20. *The isomorphism $\tilde{\psi} : \mathcal{G}(\mathcal{L}) \to \mathcal{G}(\mathcal{L}_\chi)$ (arising from Lemmas 4.12 and 4.19) is Galois equivariant. In particular, $\tilde{\psi}$ fits into a commutative diagram of G_F-modules*

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \bar{F}^\times & \longrightarrow & \mathcal{G}(\mathcal{L}) & \longrightarrow & A[2] & \longrightarrow & 0 \\ & & \| & & \downarrow{\scriptstyle \tilde{\psi}} & & \downarrow{\scriptstyle \psi} & & \\ 0 & \longrightarrow & \bar{F}^\times & \longrightarrow & \mathcal{G}(\mathcal{L}_\chi) & \longrightarrow & A^\chi[2] & \longrightarrow & 0, \end{array}$$

where all vertical maps are isomorphisms.

Proof. Write $\mathrm{Isom}_{\mathcal{L},\mathcal{L}_\chi}(A, A^\chi)$ for the set of $\bar{F}$-isomorphisms $f : A \to A^\chi$ for which $f^*\mathcal{L}_\chi \cong \mathcal{L}$. Then using the explicit Galois action given in Remark 4.14 one sees that the map $\mathrm{Isom}_{\mathcal{L},\mathcal{L}_\chi}(A, A^\chi) \to \mathrm{Isom}(\mathcal{G}(\mathcal{L}), \mathcal{G}(\mathcal{L}_\chi))$ given by $f \mapsto \tilde{f}$ is Galois equivariant. It then follows from Lemma 4.12 that we have, for all $\sigma \in G_F$,

$$(\tilde{\psi})^\sigma = \widetilde{\psi^\sigma} = \widetilde{\psi \circ [\chi(\sigma)]} = \tilde{\psi} \circ \widetilde{[\chi(\sigma)]} = \tilde{\psi}$$

where the last equality follows from Lemma 4.13. □

5. Controlling the parity of $\dim_{\mathbb{F}_2} Ш_{\mathrm{nd}}(A/K)[2]$ under quadratic twist

In this section we prove Theorem 1.4 concerning the behavior under quadratic twist of the Shafarevich–Tate group of a principally polarized abelian variety.

For the rest of the section, fix a number field K and let $(A/K, \lambda)$ be a principally polarized abelian variety. To fix notation, we briefly recall the definition of the 2-Selmer and Shafarevich–Tate groups of A/K.

5A. ***The 2-Selmer group and the Shafarevich–Tate group.*** For a place v of K we denote by $\delta_v : A(K_v)/2A(K_v) \hookrightarrow H^1(K_v, A[2])$ the connecting homomorphism associated to the multiplication-by-two Kummer sequence

$$0 \to A[2] \to A(\overline{K_v}) \xrightarrow{[2]} A(\overline{K_v}) \to 0 \tag{5.1}$$

over the completion K_v of K at v.

The 2-Selmer group of A/K is the group

$$\mathrm{Sel}_2(A/K) := \{\xi \in H^1(K, A[2]) : \xi_v \in \mathrm{im}(\delta_v)\ \forall v \in M_K\}.$$

It sits in a short exact sequence

$$0 \to A(K)/2A(K) \to \mathrm{Sel}_2(A/K) \to Ш(A/K)[2] \to 0 \tag{5.2}$$

where

$$Ш(A/K) := \ker\Big(H^1(K, A) \to \prod_{v \in M_K} H^1(K_v, A)\Big)$$

is the Shafarevich–Tate group of A/K.

5B. ***The Cassels–Tate pairing.*** Denote by $Ш_{\mathrm{nd}}(A/K)$ the quotient of $Ш(A/K)$ by its maximal divisible subgroup. The Cassels–Tate pairing is a bilinear pairing

$$\langle \cdot, \cdot \rangle_{\mathrm{CT}} : Ш(A/K) \times Ш(A^\vee/K) \to \mathbb{Q}/\mathbb{Z}$$

the left and right kernels of which are $Ш_{\mathrm{nd}}(A/K)$ and $Ш_{\mathrm{nd}}(A^\vee/K)$, respectively. The principal polarization $\lambda : A \to A^\vee$ induces a nondegenerate bilinear pairing

$$\langle \cdot, \cdot \rangle_{\mathrm{CT},\lambda} : Ш_{\mathrm{nd}}(A/K) \times Ш_{\mathrm{nd}}(A/K) \to \mathbb{Q}/\mathbb{Z}$$

defined by $\langle a, b\rangle_{\mathrm{CT},\lambda} = \langle a, \lambda(b)\rangle_{\mathrm{CT}}$ for $a, b \in Ш(A/K)$. This pairing is antisymmetric [Flach 1990, Theorem 2; Poonen and Stoll 1999, Corollary 6].

Via the map $\mathrm{Sel}_2(A/K) \to Ш(A/K)[2]$ of (5.2) the Cassels–Tate pairing $\langle \cdot, \cdot \rangle_{\mathrm{CT},\lambda}$ induces an antisymmetric pairing on $\mathrm{Sel}_2(A/K)$ (though this is no longer nondegenerate). By an abuse of notation we denote this by $\langle \cdot, \cdot \rangle_{\mathrm{CT},\lambda}$ also.

5C. ***Description of the Cassels–Tate pairing on*** $\mathbf{Sel}_2(A/K)$**.** We will need an explicit description of the Cassels–Tate pairing $\langle\cdot,\cdot\rangle_{\mathrm{CT},\lambda}$ on $\mathrm{Sel}_2(A/K)$. We use the "Weil pairing" definition as in [Poonen and Stoll 1999, §12.2] which we copy almost verbatim and to which we refer for more details.

Definition 5.3 (Cassels–Tate pairing). Let $\mathfrak{a}, \mathfrak{b} \in \mathrm{Sel}_2(A/K)$. There will be several choices involved in the definition of $\langle\mathfrak{a}, \mathfrak{b}\rangle_{\mathrm{CT},\lambda}$. We begin with the global choices.

Pick cocycles a and b representing $\mathfrak{a}$ and $\mathfrak{b}$ respectively. Next, pick $\sigma \in C^1(K, A[4])$ such that $2\sigma = a$. Then $d\sigma$ is a 2-cocycle with values in $A[2]$, i.e., an element of $Z^2(K, A[2])$. The Weil pairing $(\cdot,\cdot)_\lambda : A[2] \times A[2] \to \boldsymbol{\mu}_2 \hookrightarrow \overline{K}^\times$ induces a cup-product map $\cup : Z^2(K, A[2]) \times Z^1(K, A[2]) \to Z^3(K, \overline{K}^\times)$. As K is a number field $H^3(K, \overline{K}^\times) = 0$, so we may choose $\epsilon \in C^2(K, \overline{K}^\times)$ such that $d\sigma \cup b = d\epsilon$.

Now for the local choices. Fix a place v of K. The class of a_v is trivial in $H^1(K_v, A(\overline{K_v}))$ so we may choose $P_v \in A(\overline{K_v})$ with $a_v = dP_v$. Pick $Q_v \in A(\overline{K_v})$ with $2Q_v = P_v$. Then $\rho_v := dQ_v$ is an element of $Z^1(K_v, A[4])$ and $\sigma_v - \rho_v$ takes values in $A[2]$, i.e., is an element of $C^1(K_v, A[2])$. Then we may form the element $(\sigma_v - \rho_v) \cup b_v$ of $C^2(K_v, \overline{K_v}^\times)$ (again defining the cup-product map using the Weil pairing on $A[2]$). The difference $(\sigma_v - \rho_v) \cup b_v - \epsilon_v$ is a 2-cocycle with values in $\overline{K_v}^\times$. Let $\mathfrak{d}_v$ denote its class in $H^2(K_v, \overline{K_v}^\times) = \mathrm{Br}(K_v)$. Then $\langle\mathfrak{a}, \mathfrak{b}\rangle_{\mathrm{CT},\lambda}$ is defined as

$$\langle\mathfrak{a}, \mathfrak{b}\rangle_{\mathrm{CT},\lambda} := \sum_{v \in M_K} \mathrm{inv}_v(\mathfrak{d}_v) \in \mathbb{Q}/\mathbb{Z}.$$

The value of the sum above is independent of all choices made.

5D. ***Controlling the parity of*** $\dim_{\mathbb{F}_2} Ш_{\mathrm{nd}}(A/K)[2]$ ***globally.*** If A is an elliptic curve and λ its canonical principal polarization then it is well known that $\langle\cdot,\cdot\rangle_{\mathrm{CT},\lambda}$ is in fact alternating and it follows that $\dim_{\mathbb{F}_2} Ш_{\mathrm{nd}}(A/K)[2]$ is even. For general principally polarized abelian varieties however, Poonen and Stoll [1999] showed that $\dim_{\mathbb{F}_2} Ш_{\mathrm{nd}}(A/K)[2]$ need not be even and gave a criterion for determining whether or not this is the case. Specifically, let $\mathfrak{c}_\lambda \in H^1(K, A[2])$ be the cohomology class associated to λ as in Definition 4.6. By [Poonen and Stoll 1999, Lemma 1] we in fact have $\mathfrak{c}_\lambda \in \mathrm{Sel}_2(A/K)$.

We then have the following theorem of Poonen–Stoll.

Theorem 5.4. *The group* $Ш_{\mathrm{nd}}(A/K)[2]$ *has even* $\mathbb{F}_2$*-dimension if and only if*

$$\langle\mathfrak{c}_\lambda, \mathfrak{c}_\lambda\rangle_{\mathrm{CT},\lambda} = 0 \in \mathbb{Q}/\mathbb{Z}.$$

Proof. The image of $\mathfrak{c}_\lambda$ in $Ш(A/K)[2]$ is the homogeneous space associated to λ as in [Poonen and Stoll 1999, §2]. Theorem 8 of [loc. cit.] now gives the result. □

Remark 5.5. Since the image of $\mathfrak{c}_\lambda$ in $Ш(A/K)$ is annihilated by 2 we have $\langle\mathfrak{c}_\lambda, \mathfrak{c}_\lambda\rangle_{\mathrm{CT},\lambda} \in \left\{0, \frac{1}{2}\right\}$.

5E. ***Quadratic twists.*** For the rest of the section fix a quadratic character χ and let (A^χ, ψ) be the quadratic twist of A by χ. We now set up the notation which we will use when computing with A^χ in what follows. We endow A^χ with the K-rational principal polarization $\lambda_\chi := (\psi^\vee)^{-1}\lambda\psi^{-1}$ (see

Section 4E). Associated to λ_χ is the Weil pairing

$$(\cdot,\cdot)_{\lambda_\chi} : A^\chi[2] \times A^\chi[2] \to \boldsymbol{\mu}_2$$

and the Cassels–Tate pairing

$$\langle \cdot,\cdot \rangle_{\mathrm{CT},\lambda_\chi} : Ш(A^\chi/K)[2] \times Ш(A^\chi/K)[2] \to \mathbb{Q}/\mathbb{Z}$$

(which we also view as a pairing on $\mathrm{Sel}_2(A^\chi/K)$). Using the isomorphism ψ we identify $A^\chi[2]$ and $A[2]$ as G_K-modules. Note that this identification also respects the Weil pairing (i.e., identifies $(\cdot,\cdot)_{\lambda_\chi}$ with $(\cdot,\cdot)_\lambda$; to see this e.g., combine Lemma 4.5(ii) and Lemma 4.17). In this way, we identify $H^1(K, A^\chi[2])$ with $H^1(K, A[2])$ and thus view the 2-Selmer group $\mathrm{Sel}_2(A^\chi/K)$ inside $H^1(K, A[2])$. In particular, we may talk about the intersection of $\mathrm{Sel}_2(A/K)$ and $\mathrm{Sel}_2(A^\chi/K)$.

We also use ψ to identify $A[4](\bar{K})$ with $A^\chi[4](\bar{K})$. This last identification does not respect the G_K-action. Thus for each i, we have identified $C^i(K, A^\chi[4])$ with $C^i(K, A[4])$ but the differential $d : C^i(K, A^\chi[4]) \to C^{i+1}(K, A^\chi[4])$ is not identified with the usual differential on $C^i(K, A[4])$; we write d_χ for the map $C^i(K, A[4]) \to C^{i+1}(K, A[4])$ to which is does correspond. For example, the map $d : C^1(K, A^\chi[4]) \to C^2(K, A^\chi[4])$ corresponds to the map $d_\chi : C^1(K, A[4]) \to C^2(K, A[4])$ defined by

$$(d_\chi f)(\sigma,\tau) = f(\sigma) + \chi(\sigma)\sigma f(\tau) - f(\sigma\tau).$$

Similarly, we use ψ to identify $C^i(K, A^\chi(\bar{K}))$ and $C^i(K, A(\bar{K}))$ for each i, and define differentials d_χ on $C^i(K, A(\bar{K}))$ similarly.

5F. ***Strategy of the proof of Theorem 1.4.*** To motivate what follows, we briefly sketch the proof of Theorem 1.4.

For $\mathfrak{a}, \mathfrak{b} \in Ш(A/K)$, in the definition of $\langle \mathfrak{a}, \mathfrak{b} \rangle_{\mathrm{CT},\lambda}$ the local terms $\mathfrak{d}_v$ (in the notation of Definition 5.3) depend on the global choices. In particular, it is not clear that $\langle \mathfrak{c}_\lambda, \mathfrak{c}_\lambda \rangle_{\mathrm{CT},\lambda}$, and hence the parity of $\dim_{\mathbb{F}_2} Ш_{\mathrm{nd}}(A/K)[2]$, may be expressed as a sum of local terms whose definition requires no global choices (this is, however, known to be true if A/K is the Jacobian of a curve, see [Poonen and Stoll 1999, Corollary 12]).

When considering A along with its quadratic twist A^χ, we eliminate the global choices as follows. Associated to λ_χ is the class $\mathfrak{c}_{\lambda_\chi} \in \mathrm{Sel}_2(A^\chi/K)$ (viewed inside $H^1(K, A[2])$ as in Section 5E). By Corollary 4.18 we have $\mathfrak{c}_{\lambda_\chi} = \mathfrak{c}_\lambda$ and in particular, $\mathfrak{c}_\lambda$ lies in $\mathrm{Sel}_2(A/K) \cap \mathrm{Sel}_2(A^\chi/K)$. Now the sum of the pairings $\langle \cdot,\cdot \rangle_{\mathrm{CT},\lambda}$ and $\langle \cdot,\cdot \rangle_{\mathrm{CT},\lambda_\chi}$ gives a new pairing on $\mathrm{Sel}_2(A/K) \cap \mathrm{Sel}_2(A^\chi/K)$. By Theorem 5.4, $\dim_{\mathbb{F}_2} Ш_{\mathrm{nd}}(A/K)[2] + \dim_{\mathbb{F}_2} Ш_{\mathrm{nd}}(A^\chi/K)[2]$ is even if and only if $\mathfrak{c}_\lambda$ pairs trivially with itself under this new pairing.

We show in Lemma 5.8 that the global choices involved in computing the sum of the two Cassels–Tate pairings are milder than those for the individual pairings (we remark that this simplification of the Cassels–Tate pairing under quadratic twist has also been observed in the recent preprint of Smith [2016,

proof of Theorem 3.2]). Specifically, the global choices involved in computing

$$\langle \mathfrak{c}_\lambda, \mathfrak{c}_\lambda\rangle_{\mathrm{CT},\lambda} + \langle \mathfrak{c}_\lambda, \mathfrak{c}_\lambda\rangle_{\mathrm{CT},\lambda_\chi}$$

are: a choice of cocycle $c_\lambda \in Z^1(K, A[2])$ representing $\mathfrak{c}_\lambda$ and a choice of cochain $F : G_K \to \boldsymbol{\mu}_2$ such that $dF = c_\lambda \cup c_\lambda \in Z^2(K, \boldsymbol{\mu}_2)$.

By Lemma 4.7, $\mathfrak{c}_\lambda \in H^1(K, A[2])$ is the cohomology class parametrizing quadratic refinements of the Weil pairing. In particular, a choice of cocycle representing $\mathfrak{c}_\lambda$ amounts to a choice of quadratic refinement q. For each such q we have already constructed a canonical choice for the function F above, namely that given by Proposition 3.9. Thus the only global choice remaining is that of q. Proposition 3.10 shows how this choice for F changes upon changing q, allowing us to prove that the local terms then arising do not, in fact, depend on the choice of quadratic refinement either.

5G. ***Pairings on*** $\mathbf{Sel_2}(A/K) \cap \mathbf{Sel_2}(A^\chi/K)$**.** Define $\mathcal{S}_\chi := \mathrm{Sel}_2(A/K) \cap \mathrm{Sel}_2(A^\chi/K)$. Here we define a pairing $\langle \cdot, \cdot \rangle_{S_\chi}$ on $\mathcal{S}_\chi$ with values in $\mathbb{Q}/\mathbb{Z}$ which we shall see is the sum of the Cassels–Tate pairings for A and its twist A^χ. However, for clarity when using this pairing later, we define it separately.

Definition 5.6 (the pairing $\langle \cdot, \cdot \rangle_{S_\chi}$). Let $\mathfrak{a},\mathfrak{b} \in \mathcal{S}_\chi = \mathrm{Sel}_2(A/K) \cap \mathrm{Sel}_2(A^\chi/K)$. As with the definition of the Cassels–Tate pairing, we begin with the global choices. We first claim that $\mathfrak{a} \cup \mathfrak{b} = 0 \in H^2(K, \boldsymbol{\mu}_2) = \mathrm{Br}(K)[2]$. Indeed, for each place v of K both $\mathfrak{a}_v$ and $\mathfrak{b}_v$ are in the image of $A(K_v)/2A(K_v)$ under the connecting homomorphism associated to the multiplication-by-2 Kummer sequence. Since this image is its own orthogonal complement under the cup-product pairing

$$H^1(K_v, A[2]) \times H^1(K_v, A[2]) \to H^2(K_v, \overline{K_v}^\times) = \mathrm{Br}(K_v)$$

(this results from Tate local duality, see e.g., [Milne 2006, I.3.4]) we have $(\mathfrak{a} \cup \mathfrak{b})_v = 0 \in \mathrm{Br}(K_v)$ for each place v of K. Reciprocity for the Brauer group now gives the claim.

Now represent $\mathfrak{a}$ and $\mathfrak{b}$ by cocycles a and b respectively and, as is possible by the above discussion, pick $f \in C^1(K, \boldsymbol{\mu}_2)$ with $df = a \cup b \in Z^2(K, \boldsymbol{\mu}_2)$.

We now turn to the local choices. Fix a place v of K. Since $\mathfrak{a} \in \mathrm{Sel}_2(A/K)$ there is $P_v \in A(\overline{K_v})$ with $dP_v = a_v$. Pick $Q_v \in A(\overline{K_v})$ with $2Q_v = P_v$. Then $\rho_v := dQ_v$ is an element of $Z^1(K_v, A[4])$. Since $\mathfrak{a}$ is also in $\mathrm{Sel}_2(A^\chi/K)$ we can similarly (i.e., by replacing d by d_χ throughout) define $P_{v,\chi}$, $Q_{v,\chi}$ and $\rho_{v,\chi} = d_\chi Q_{v,\chi} \in C^1(K_v, A[4])$. Then $\rho_v + \rho_{v,\chi}$ takes values in $A[2]$. One checks that $d(\rho_v + \rho_{v,\chi}) = \chi_v \cup a_v \in Z^2(K_v, A[2])$. Thus the difference

$$(\rho_v + \rho_{v,\chi}) \cup b_v - \chi_v \cup f_v$$

is a 2-cocycle with values in $\boldsymbol{\mu}_2$. Denote by $\mathfrak{d}_v$ its class in $\mathrm{Br}(K_v)[2]$.

Now define

$$\langle a, b\rangle_{S_\chi} := \sum_{v \in M_K} \mathrm{inv}_v(\mathfrak{d}_v) \in \mathbb{Q}/\mathbb{Z}.$$

One easily checks that once the initial global choices are made the cocycle class $\mathfrak{d}_v \in \mathrm{Br}(K_v)$ is independent of the local choices. That the resulting sum is independent of all choices follows from reciprocity for the Brauer group.

Remark 5.7. If a place v of K splits in the quadratic extension L/K associated to χ then χ_v is trivial and ψ gives an isomorphism between A and A^χ over K_v. It follows easily that the local terms $\mathrm{inv}_v(\mathfrak{d}_v)$ are trivial at all such v. Thus in the definition of $\langle \cdot, \cdot \rangle_{S_\chi}$ we may replace the sum over all places of K by the sum over all places of K nonsplit in L/K.

Lemma 5.8. *The pairing $\langle \cdot, \cdot \rangle_{S_\chi}$ is the sum of the Cassels–Tate pairings for A and A^χ:*

$$\langle \cdot, \cdot \rangle_{S_\chi} = \langle \cdot, \cdot \rangle_{CT,\lambda} + \langle \cdot, \cdot \rangle_{CT,\lambda_\chi}.$$

In particular, it is (anti)symmetric.

Remark 5.9. This lemma is implicit in the recent preprint of Smith [2016, proof of Theorem 3.2].

Proof. Fix $\mathfrak{a}, \mathfrak{b} \in S_\chi$. We begin by making the global choices involved in computing $\langle \mathfrak{a}, \mathfrak{b} \rangle_{CT,\lambda}$. We pick cocycles a and b representing $\mathfrak{a}$ and $\mathfrak{b}$ respectively and pick $\sigma \in C^1(K, A[4])$ with $2\sigma = a$. Next, we pick $\epsilon \in C^2(K, \overline{K}^\times)$ with $d\epsilon = d\sigma \cup b$.

We now make the corresponding choices involved in computing $\langle \mathfrak{a}, \mathfrak{b} \rangle_{CT,\lambda_\chi}$. As we are at liberty to do, we pick the same cocycle representatives a and b chosen above. We similarly pick the same element σ of $C^1(K, A[4])$ satisfying $2\sigma = a$ (here using the identification of $A[4]$ with $A^\chi[4]$ via ψ as discussed). We then pick $\epsilon_\chi \in C^2(K, \overline{K}^\times)$ such that $d\epsilon_\chi = d_\chi\sigma \cup b$. Note that we cannot chose $\epsilon = \epsilon_\chi$ in general due to the difference between the differentials d and d_χ. However, we have

$$d(\epsilon + \epsilon_\chi) = (d\sigma + d_\chi\sigma) \cup b = (\chi \cup a) \cup b,$$

the last equality following from the definition of d_χ and a simple computation.

Now let $f \in C^1(K, \boldsymbol{\mu}_2)$ be such that $df = a \cup b$. By (2.1) and associativity of the cup-product we have $d(\chi \cup f) = d(\epsilon + \epsilon_\chi)$ whence $\chi \cup f = \epsilon + \epsilon_\chi + \nu$ for some cocycle $\nu \in Z^2(K, \overline{K}^\times)$.

We now make the local choices involved in computing $\langle \mathfrak{a}, \mathfrak{b} \rangle_{CT,\lambda}$. We choose $P_v \in A(\overline{K_v})$ with $dP_v = a_v$ and then pick $Q_v \in A(\overline{K_v})$ with $2Q_v = P_v$. Next, set $\rho_v := dQ_v \in C^1(K_v, A[4])$ and define $\mathfrak{d}_v$ to be the class of $(\sigma_v - \rho_v) \cup b_v - \epsilon_v$ in $H^2(K_v, \overline{K_v}^\times)$.

Finally, we make the local choices involved in computing $\langle \mathfrak{a}, \mathfrak{b} \rangle_{CT,\lambda_\chi}$. Thus we pick $P_{v,\chi}$ with $d_\chi P_{v,\chi} = a_v$, $Q_{v,\chi}$ with $2Q_{v,\chi} = P_{v,\chi}$, set $\rho_{v,\chi} = d_\chi Q_{v,\chi}$ and define $\mathfrak{d}_{v,\chi}$ to be the class of $(\sigma_v - \rho_{v,\chi}) \cup b_v - \epsilon_{\chi,v}$ in $H^2(K_v, \overline{K_v}^\times)$.

With these choices in place $\mathfrak{d}_v + \mathfrak{d}_{v,\chi}$ is the class in $\mathrm{Br}(K_v)$ of

$$(a_v - (\rho_v + \rho_{v,\chi})) \cup b_v - \chi_v \cup f_v + \nu_v.$$

Noting that $\rho_v + \rho_{v,\chi}$ takes values in $A[2]$ and that $\mathfrak{a}_v \cup \mathfrak{b}_v = 0$ (as discussed previously) we see that

$$\mathrm{inv}_v(\mathfrak{d}_v) + \mathrm{inv}_v(\mathfrak{d}_{v,\chi}) = \mathrm{inv}_v((\rho_v + \rho_{v,\chi}) \cup b_v - \chi_v \cup f_v) + \mathrm{inv}_v(\nu_v).$$

Summing over all places and noting that by reciprocity for the Brauer group we have

$$\sum_{v \in M_K} \mathrm{inv}_v(\nu_v) = 0 \in \mathbb{Q}/\mathbb{Z},$$

we have

$$\langle \mathfrak{a}, \mathfrak{b} \rangle_{CT} + \langle \mathfrak{a}, \mathfrak{b} \rangle_{CT,\chi} = \sum_{v \in M_K} \mathrm{inv}_v((\rho_v + \rho_{v,\chi}) \cup b_v - \chi_v \cup f_v).$$

But this is precisely how the quantity $\langle \mathfrak{a}, \mathfrak{b} \rangle_{S_\chi}$ was defined. □

5H. ***The local terms*** $\mathfrak{g}(A, \lambda, \chi)$**.** In this subsection we study the local terms which arise in computing $\langle \mathfrak{c}_\lambda, \mathfrak{c}_\lambda \rangle_{\mathcal{S}_\chi}$, and show in particular that they are independent of certain choices involved. We work purely locally and take F to be a local field of characteristic 0. Let $(A/F, \lambda)$ be a principally polarized abelian variety. Let $\chi \in \mathrm{Hom}_{\mathrm{cnt}}(G_F, \boldsymbol{\mu}_2)$ be a quadratic character of F and $(A^\chi/F, \psi)$ be the quadratic twist of A by χ. We use the same conventions and notation as in Section 5E when talking about objects associated to A^χ. We will need to identify $\boldsymbol{\mu}_2$ with the additive group of $\mathbb{F}_2$ in the following, and we write the group law on $\boldsymbol{\mu}_2$ additively to avoid confusion when doing this.

Denote by $\mathfrak{c}_\lambda \in H^1(F, A[2])$ the cohomology class associated to λ as in Definition 4.6. By [Poonen and Stoll 1999, Lemma 1] its image in $H^1(F, A)[2]$ is trivial. By Corollary 4.18, it follows also that the image of $\mathfrak{c}_\lambda$ in $H^1(F, A^\chi)[2]$ is trivial too (here the map $H^1(F, A[2]) \to H^1(F, A^\chi)[2]$ comes from identifying $A[2]$ with $A^\chi[2]$ via ψ).

Remark 5.10. By Lemma 4.7 $\mathfrak{c}_\lambda$ is equal to the cohomology class associated to the set of quadratic refinements of the Weil pairing $(\cdot, \cdot)_\lambda$ on $A[2]$. In particular, for each quadratic refinement q of $(\cdot, \cdot)_\lambda$, the function $c_q : G_F \to A[2]$ sending $\sigma \in G_F$ to the unique element $c_q(\sigma) \in A[2]$ such that

$$q(\sigma^{-1} v) - q(v) = (v, c_q(\sigma))_\lambda$$

for all $v \in A[2]$, is a cocycle in $Z^1(F, A[2])$ representing the class $\mathfrak{c}_\lambda$.

Definition 5.11. Let $q : A[2] \to \boldsymbol{\mu}_2$ be a quadratic refinement of the Weil pairing $(\cdot, \cdot)_\lambda$. Then we define the function $F_q : G_F \to \boldsymbol{\mu}_2$ as the composition

$$F_q : G_F \to \mathrm{Sp}(A[2]) \xrightarrow{f_q} \mathbb{F}_2 \cong \boldsymbol{\mu}_2,$$

where the map $G_F \to \mathrm{Sp}(A[2])$ is the homomorphism coming from the action of G_F on $A[2]$ and $f_q : \mathrm{Sp}(A[2]) \to \mathbb{F}_2$ is the map afforded by Proposition 3.9.

Remark 5.12. For each quadratic refinement q of $(\cdot, \cdot)_\lambda$ it follows from Proposition 3.9 that we have $dF_q = c_q \cup c_q \in Z^2(F, \boldsymbol{\mu}_2)$.

Definition 5.13. Let $\chi \in \mathrm{Hom}_{\mathrm{cnt}}(G_F, \boldsymbol{\mu}_2)$ be a quadratic character, let q be a quadratic refinement of $(\cdot, \cdot)_\lambda$ and let c_q be the associated cocycle representing $\mathfrak{c}_\lambda$. As in the definition of the local choices for the pairing $\langle \cdot, \cdot \rangle_{S_\chi}$, pick $P_q \in A(\bar{F})$ with $dP_q = c_q$, let $Q_q \in A(\bar{F})$ be such that $2Q_q = P_q$ and set

$\rho_q = dQ_q$. Similarly, pick $P_{\chi,q} \in A(\bar{F})$ with $d_\chi P_{\chi,q} = c_q$, let $Q_{\chi,q} \in A(\bar{F})$ be such that $2Q_{\chi,q} = P_{\chi,q}$ and set $\rho_{\chi,q} = d_\chi Q_{\chi,q}$.

We then define $\mathfrak{g}(A, \lambda, \chi, q)$ to be the class of the cocycle

$$g(A, \lambda, \chi, q) := (\rho_q + \rho_{\chi,q}) \cup c_q - \chi \cup F_q$$

in $\mathrm{Br}(F)[2]$. As in Section 5G, $\mathfrak{g}(A, \lambda, \chi, q)$ does not depend on the choices of P_q, Q_q, $P_{\chi,q}$ or $Q_{\chi,q}$.

The following lemma is key to the proof of Theorem 1.4.

Lemma 5.14. *The quantity $\mathfrak{g}(A, \lambda, \chi, q) \in \mathrm{Br}(F)[2]$ is independent of the choice of quadratic refinement q.*

Proof. Keep the notation of Definition 5.13 in what follows. Let q and q' be two quadratic refinements. Then $q - q' = (-, v)_\lambda$ for some $v \in A[2]$ and $c_{q'} = c_q + dv$. By Proposition 3.10 we have

$$F_{q'} = F_q + c_q \cup v + v \cup c_q + v \cup dv.$$

Now fix choices for P_q, Q_q, $P_{\chi,q}$ and $Q_{\chi,q}$ as in Definition 5.13. Then we may take $P_{q'} = P_q + v$ and $P_{\chi,q'} = P_{\chi,q} + v$. Pick $T \in A[4]$ with $2T = v$. Then we may take $Q_{q'} = Q_q + T$ and $Q_{\chi,q'} = Q_{\chi,q} + T$. Thus

$$\rho_{q'} + \rho_{\chi,q'} = \rho_q + \rho_{\chi,q} + dT + d_\chi T.$$

An easy computation gives $dT + d_\chi T = dv + \chi \cup v$. Combining this with the expressions for $F_{q'}$ and $c_{q'}$ in terms of F_q and c_q respectively, we see that we have an equality of cocycles

$$g(A, \lambda, \chi, q') = g(A, \lambda, \chi, q) + (\rho_q + \rho_{\chi,q}) \cup dv + (dv + \chi \cup v) \cup (c_q + dv) - \chi \cup (c_q \cup v + v \cup c_q + v \cup dv)$$

inside $Z^2(F, \boldsymbol{\mu}_2)$.

Now $c_q + dv \in C^1(F, A[2])$ is a cocycle whilst $dv \in C^1(F, A[2])$ is a coboundary. Thus the class of $dv \cup (c_q + dv)$ is trivial in $\mathrm{Br}(F)[2]$. Using this observation, canceling like terms in the previous expression, and passing to classes in the Brauer group, one has

$$\mathfrak{g}(A, \lambda, \chi, q') = \mathfrak{g}(A, \lambda, \chi, q) + [(\rho_q + \rho_{\chi,q}) \cup dv - \chi \cup c_q \cup v]$$

(where here "[ ]" denotes the operation of taking classes in the Brauer group).

Now, as remarked in the definition of the pairing $\langle \cdot, \cdot \rangle_{S_\chi}$, we have $d(\rho_q + \rho_{\chi,q}) = \chi \cup c_q$. Thus by standard properties of cup product on cochains (see Section 2B) we have

$$d((\rho_q + \rho_{\chi,q}) \cup v) = (\rho_q + \rho_{\chi,q}) \cup dv - \chi \cup c_q \cup v.$$

In particular, the class of $(\rho_q + \rho_{\chi,q}) \cup dv - \chi \cup c_q \cup v$ is trivial in $\mathrm{Br}(F)$, whence $\mathfrak{g}(A, \lambda, \chi, q') = \mathfrak{g}(A, \lambda, \chi, q)$ as desired. □

Lemma 5.14 allows us to make the following refinement of Definition 5.13.

Definition 5.15. Define $\mathfrak{g}(A, \lambda, \chi) \in \mathrm{Br}(F)[2]$ to be the quantity $\mathfrak{g}(A, \lambda, \chi, q)$ for any choice of quadratic refinement q of $(\cdot, \cdot)_\lambda$.

The following proposition computes explicitly the terms $\mathfrak{g}(A, \lambda, \chi)$ in certain cases.

Proposition 5.16. *Let $\mathfrak{g}(A, \lambda, \chi) \in \mathrm{Br}(F)[2]$ be as in Definition 5.15.*

(i) *We have $\mathfrak{g}(A, \lambda, \mathbb{1}) = 0$ where $\mathbb{1}$ is the trivial character of F.*

(ii) *Suppose that q is a G_F-invariant quadratic refinement of the Weil pairing $(\cdot, \cdot)_\lambda$ on $A[2]$ and let $\alpha : G_F \to \boldsymbol{\mu}_2$ be the quadratic character corresponding to the homomorphism*

$$G_F \to O(q)/\mathrm{SO}(q) \cong \mathbb{Z}/2\mathbb{Z} \cong \boldsymbol{\mu}_2$$

coming from the action of G_F on $A[2]$. Then

$$\mathfrak{g}(A, \lambda, \chi) = \alpha \cup \chi \in \mathrm{Br}(F)[2].$$

(iii) *Suppose that F is nonarchimedean with odd residue characteristic and that A has good reduction. Then we have*

$$\mathrm{inv}_F\, \mathfrak{g}(A, \lambda, \chi) = \begin{cases} 0 & \chi \text{ unramified}, \\ \frac{1}{2} \dim_{\mathbb{F}_2} A(F)[2] \in \mathbb{Q}/\mathbb{Z} & \chi \text{ ramified}. \end{cases}$$

(iv) *Suppose that F is archimedean. Then we have*

$$\mathrm{inv}_F\, \mathfrak{g}(A, \lambda, \chi) = \begin{cases} 0 & F = \mathbb{C} \text{ or } \chi \text{ trivial}, \\ \frac{1}{2} \dim_{\mathbb{F}_2} A(F)[2] \in \mathbb{Q}/\mathbb{Z} & F = \mathbb{R} \text{ and } \chi \text{ nontrivial}. \end{cases}$$

Proof.

(i) Clear.

(ii) If there is an F-rational quadratic refinement q then c_q is identically zero and it follows immediately from Lemma 5.14 and the definition of $\mathfrak{g}(A, \lambda, \chi, q)$ that

$$\mathfrak{g}(A, \lambda, \chi) = \mathfrak{g}(A, \lambda, \chi, q) = -\chi \cup F_q = \chi \cup \alpha$$

where for the last equality we use that the restriction of F_q to elements of $O(q)$ agrees with the Dickson homomorphism d_q (see Proposition 3.9).

(iii) By [Poonen and Rains 2011, Proposition 3.6(d)], our assumptions on F and the reduction of A imply that there is a G_F-invariant quadratic refinement q of the Weil pairing on $A[2]$. Let α be the associated quadratic character so that $\mathfrak{g}(A, \lambda, \chi) = \alpha \cup \chi$ by (ii). Now by definition, α factors through $\mathrm{Gal}(F(A[2])/F)$ and our assumptions on F and A mean that $F(A[2])/F$ is unramified. Consequently, α is unramified. In fact, let σ denote the Frobenius element in $F(A[2])/F$. Then by Proposition 3.5 we have

$$\alpha(\sigma) = (-1)^{\dim_{\mathbb{F}_2} A[2]^\sigma} = (-1)^{\dim_{\mathbb{F}_2} A(F)[2]}.$$

In particular, we see that if $\dim_{\mathbb{F}_2} A(F)[2]$ is even then α is the trivial character, whilst if $\dim_{\mathbb{F}_2} A(F)[2]$ is odd then α is the unique nontrivial unramified quadratic character of F. Since F is assumed to have odd residue characteristic, standard properties of the cup-product of two quadratic characters gives the result (we review these later in Section 8A: see, in particular, Lemma 8.4).

(iv) The argument here is similar to that of (iii). First note that if χ is trivial then $\mathfrak{g}(A, \lambda, \chi) = 0$ by (i). In particular, the only case we have not already covered is when $F = \mathbb{R}$ and χ is the quadratic character corresponding to the extension $\mathbb{C}/\mathbb{R}$. By [Poonen and Rains 2011, Proposition 3.6(d)] there is an $\mathbb{R}$-rational quadratic refinement q of the Weil pairing $(\cdot, \cdot)_\lambda$. Let α be the associated quadratic character and write σ for the unique nontrivial element of $\mathrm{Gal}(\mathbb{C}/\mathbb{R})$. By Proposition 3.5 we see that α is trivial if $\dim_{\mathbb{F}_2} A[2]^\sigma = \dim_{\mathbb{F}_2} A(\mathbb{R})[2]$ is even, and is the quadratic character corresponding to $\mathbb{C}/\mathbb{R}$ otherwise. The result now follows from (ii). □

Remark 5.17. As in Lemma 4.5, if the polarization λ is of the form $\phi_{\mathcal{L}}$ for an F-rational symmetric line bundle $\mathcal{L}$ then there is an associated G_F-invariant quadratic refinement of the Weil pairing on $A[2]$. Thus combined with Proposition 5.16(ii) this gives a geometric condition for when the local terms $\mathfrak{g}(A, \lambda, \chi)$ may be evaluated.

Remark 5.18. It is natural to ask if the terms $\mathfrak{g}(A, \lambda, \chi)$ are independent of the choice of principal polarization λ. The above proposition shows that this is true when χ is trivial, when A/F has good reduction and F has odd residue characteristic, or when F is archimedean. We have been unable to prove this in general however.

Remark 5.19. Write $L = F(A[2])$ and let χ be any quadratic character. Since for any quadratic refinement q the cocycle c_q factors through $\mathrm{Gal}(L/F)$, the points P_q and $P_{\chi.q}$ of Definition 5.13 lie in $A(L)$ and $A^\chi(L)$ respectively. In particular, it follows that the cocycle $\mathfrak{g}(A, \lambda, \chi)$ factors through $\mathrm{Gal}(L'/F)$, where L' is the compositum of all the (finitely many) quadratic extensions of $F(A[2])$.

5I. ***Controlling the parity of*** $\dim_{\mathbb{F}_2} Ш_{\mathrm{nd}}(A/K)[2] + \dim_{\mathbb{F}_2} Ш_{\mathrm{nd}}(A^\chi/K)[2]$ ***via local contributions.*** We return to the notation of Section 5A–5G so that, in particular, K is a number field and $(A/K, \lambda)$ a principally polarized abelian variety.

Theorem 5.20 (Theorem 1.4). *Let χ be a quadratic character of K and for each place v of K write χ_v for the restriction of χ to G_{K_v}, A/K_v for the base change of A to K_v, and λ_v for the principal polarization on A/K_v corresponding to λ.*

Then $\dim_{\mathbb{F}_2} Ш_{\mathrm{nd}}(A/K)[2] + \dim_{\mathbb{F}_2} Ш_{\mathrm{nd}}(A^\chi/K)[2] \equiv 0 \pmod 2$ *if and only if*

$$\sum_{v \in M_K} \mathrm{inv}_v\, \mathfrak{g}(A/K_v, \lambda_v, \chi_v) = 0 \in \mathbb{Q}/\mathbb{Z}.$$

Remark 5.21. Before proving Theorem 5.20 we remark that if v is a nonarchimedean place of K, not dividing 2 and such that both A has good reduction and χ is unramified at v, then $\mathfrak{g}(A/K_v, \lambda_v, \chi_v) = 0$ by Proposition 5.16(iii). In particular, the sum in the statement of Theorem 5.20 is finite.

Proof of Theorem 5.20. By Corollary 4.18, Theorem 5.4 applied to both A and A^χ (along with their principal polarizations λ and λ_χ), and Lemma 5.8, we see that $\dim_{\mathbb{F}_2} Ш_{\mathrm{nd}}(A/K)[2] + \dim_{\mathbb{F}_2} Ш_{\mathrm{nd}}(A_\chi/K)[2]$ is even if and only if $\langle \mathfrak{c}_\lambda, \mathfrak{c}_\lambda \rangle_{\mathcal{S}_\chi} = 0$.

We now follow Definition 5.6 to compute $\langle \mathfrak{c}_\lambda, \mathfrak{c}_\lambda \rangle_{\mathcal{S}_\chi}$. For the global choices, fix a quadratic refinement q of the Weil pairing $(\cdot, \cdot)_\lambda$ on $A[2]$. Then as in the local case (Remark 5.10) the function $c_q : G_K \to A[2]$ sending $\sigma \in G_K$ to the unique element $c_q(\sigma) \in A[2]$ such that

$$q(\sigma^{-1} v) - q(v) = (v, c_q(\sigma))_\lambda$$

for all $v \in A[2]$, is a cocycle in $Z^1(F, A[2])$ representing the class $\mathfrak{c}_\lambda$. Similarly, the function $F_q : G_K \to \boldsymbol{\mu}_2$ defined as the composition

$$F_q : G_K \to \mathrm{Sp}(A[2]) \xrightarrow{f_q} \mathbb{F}_2 \cong \boldsymbol{\mu}_2,$$

(where the map $G_K \to \mathrm{Sp}(A[2])$ is the homomorphism coming from the action of G_K on $A[2]$ and $f_q : \mathrm{Sp}(A[2]) \to \mathbb{F}_2$ is the map afforded by Proposition 3.9) is an element of $C^1(K, \boldsymbol{\mu}_2)$ satisfying $dF_q = c_q \cup c_q \in Z^2(K, \bar{K}^\times)$.

With these global choices in place, the local terms arising in the definition of $\langle \mathfrak{c}_\lambda, \mathfrak{c}_\lambda \rangle_{\mathcal{S}_\chi}$ are precisely the terms $\mathfrak{g}(A/K_v, \lambda_v, \chi_v, q)$ of Definition 5.13. By Lemma 5.14 (for fixed v) they are independent of q, their common value being by definition $\mathfrak{g}(A/K_v, \lambda_v, \chi_v)$.

Thus

$$\langle \mathfrak{c}_\lambda, \mathfrak{c}_\lambda \rangle_{\mathcal{S}_\chi} = \sum_{v \in M_K} \mathrm{inv}_v\, \mathfrak{g}(A/K_v, \lambda_v, \chi_v)$$

and the result follows. □

6. Disparity in Selmer ranks: definitions and recollections

The next four sections are devoted to proving Theorem 7.4 concerning the parity of certain Selmer groups defined in terms of abstract twisting data. Our approach follows closely the strategy of [Klagsbrun et al. 2013], which proves the result for Galois modules of dimension 2 (whilst we handle arbitrary (even) dimension). Many of the statements of [loc. cit.] go through with some minor changes however in order to highlight the differences it is necessary to recall much of their setup and basic results. Thus in this section we recall the setup of [loc. cit.]. Where notions need to be generalized or slightly adapted we state the differences in a remark immediately following the definition.

6A. ***Notation.*** Here we fix some notation which will remain in place for the entirety of Sections 6–9. Fix first a prime p and number field K. Following [loc. cit.], for a field L (either K or K_v for some $v \in M_K$) we define $\mathcal{C}(L) := \mathrm{Hom}_{\mathrm{cnt}}(G_L, \boldsymbol{\mu}_p)$, the group of characters of order dividing p. We denote the trivial character by $\mathbb{1}_L$. Further, we define $\mathcal{F}(L)$ to be the quotient of $\mathcal{C}(L)$ by the action of $\mathrm{Aut}(\boldsymbol{\mu}_p)$ (the action given by post-composition). The set $\mathcal{F}(L)$ is naturally identified with the set of cyclic extensions of L of degree dividing p, the map being given by sending the equivalence class of $\chi \in \mathcal{C}(L)$ to the fixed field

$\overline{K}^{\ker(\chi)}$. When L is a nonarchimedean local field we write $\mathcal{C}_{\mathrm{ram}}(L)$ and $\mathcal{C}_{\mathrm{ur}}(L)$ for the subsets of $\mathcal{C}(L)$ consisting of ramified and unramified characters, and similarly write $\mathcal{F}_{\mathrm{ram}}(L)$ and $\mathcal{F}_{\mathrm{ur}}(L)$ for the subsets of $\mathcal{F}(L)$ corresponding to ramified and unramified extensions. Note that if L has residue characteristic coprime to p then $\mathcal{C}_{\mathrm{ram}}(L)$ (and hence also $\mathcal{F}_{\mathrm{ram}}(L)$) is nonempty if and only if $\boldsymbol{\mu}_p \subseteq L$.

For an finite dimensional $\mathbb{F}_p$-vector space M we say that a map $q : M \to \mathbb{Q}/\mathbb{Z}$ is a *quadratic form* if $q(nx) = n^2 q(x)$ for all $n \in \mathbb{Z}$ and $x \in M$, and if the map $(x, y) \mapsto q(x+y) - q(x) - q(y)$ is a symmetric bilinear pairing on M. We say that q is *nondegenerate* if the associated pairing is (i.e., if it has trivial kernel). If q is a quadratic form on M with associated pairing $\langle \cdot, \cdot \rangle$ then for a subspace W of M we write

$$W^{\perp} = \{m \in M : \langle w, m \rangle = 0, \ \forall w \in W\}$$

for the orthogonal complement of W and say that W is a *Lagrangian* subspace of (M, q) if $W = W^{\perp}$ and $q(W) = 0$. We call (M, q) a *metabolic space* if q is nondegenerate and if M has a Lagrangian subspace.

6B. ***The module T and the finite set of places Σ.*** Fix, for the remainder of Sections 6–9, a finite dimensional $\mathbb{F}_p$-vector space T equipped with a continuous G_K-action and a nondegenerate G_K-equivariant alternating pairing

$$(\cdot, \cdot) : T \times T \to \boldsymbol{\mu}_p$$

(so that, in particular, $\dim_{\mathbb{F}_p} T$ is necessarily even). For $v \in M_K$, if the inertia subgroup of G_{K_v} acts trivially on T then we say that T is *unramified* at v, and *ramified* at v otherwise. We denote by $K(T)$ the field of definition of the elements of T, i.e., the fixed field of the kernel of the action of G_K on T. Note that the presence of the pairing forces $K(\boldsymbol{\mu}_p) \subseteq K(T)$.

We also fix a finite set Σ of places of K containing all archimedean places, all places over p, and all places where T is ramified (and possibly some more to be specified later).

6C. ***The local Tate pairing and Tate quadratic forms.*** For each place $v \in M_K$ write $\langle \cdot, \cdot \rangle_v$ for the local Tate pairing

$$H^1(K_v, T) \times H^1(K_v, T) \to \mathbb{Q}/\mathbb{Z}$$

given by the composition

$$H^1(K_v, T) \times H^1(K_v, T) \xrightarrow{\cup} H^2(K_v, \boldsymbol{\mu}_p) \xrightarrow{\mathrm{inv}_v} \mathbb{Q}/\mathbb{Z},$$

where the first map is induced by cup-product and the pairing $(\cdot, \cdot)$. It is nondegenerate, bilinear and symmetric.

Definition 6.1. Let v be a place of K. We say a quadratic form $q_v : H^1(K_v, T) \to \mathbb{Q}/\mathbb{Z}$ is a *Tate quadratic form* if its associated bilinear form is the local Tate pairing $\langle \cdot, \cdot \rangle_v$. If $v \notin \Sigma$ then we say that q_v is *unramified* if it vanishes on $H^1_{\mathrm{ur}}(K_v, T)$ (in which case $H^1_{\mathrm{ur}}(K_v, T)$ is a Lagrangian subspace for q_v).

Remark 6.2. If $p = 2$ then our definition differs slightly from that of Klagsbrun, Mazur and Rubin [2013, Definition 3.2] since it allows quadratic forms valued in $\frac{1}{4}\mathbb{Z}/\mathbb{Z}$ whilst their definition only allows quadratic

forms taking values in $\frac{1}{2}\mathbb{Z}/\mathbb{Z}$. This extra generality is necessary when $\dim_{\mathbb{F}_p} T > 2$ in order to allow $T = A[2]$ for a principally polarized abelian variety A/K (see Remark 10.4).

As in [loc. cit., Lemma 3.4], if $p > 2$ then there is a unique Tate quadratic form q_v on $H^1(K_v, T)$ given by

$$q_v = \tfrac{1}{2}\langle \cdot, \cdot \rangle_v.$$

6D. ***Global metabolic structures.*** With our slightly modified definition of a Tate quadratic form in hand we can define a global metabolic structure on T in an identical way to [loc. cit., Definition 3.3].

Definition 6.3. A *global metabolic structure* $\boldsymbol{q}$ on T consists of a collection $\boldsymbol{q} = (q_v)_v$ $(v \in M_K)$ of Tate quadratic forms such that:

(i) For each $v \in M_K$ the pair $(H^1(K_v, T), q_v)$ is a metabolic space.

(ii) The quadratic form q_v is unramified at each place $v \notin \Sigma$.

(iii) If $c \in H^1(K, T)$ then $\sum_v q_v(c_v) = 0$.

As in [loc. cit., Lemma 3.4], if $p > 2$ then the unique Tate quadratic forms on $H^1(K_v, T)$ defined above do indeed give a global metabolic structure on T, so specifying a global metabolic structure is only necessary when $p = 2$.

6E. ***Selmer structures and Selmer groups.*** We define Selmer structures for $(T, \boldsymbol{q})$, along with the associated Selmer groups, as in [loc. cit., Definition 3.8].

Definition 6.4. A Selmer structure $\mathcal{S}$ for $(T, \boldsymbol{q})$ is the data:

(i) A finite set $\Sigma_{\mathcal{S}}$ of places of K containing Σ.

(ii) For each $v \in \Sigma_{\mathcal{S}}$ a Lagrangian subspace $H_{\mathcal{S}}(K_v, T)$ of $(H^1(K_v, T), q_v)$.

Definition 6.5. Let $\mathcal{S}$ be a Selmer structure for $(T, \boldsymbol{q})$. For each $v \notin \Sigma_{\mathcal{S}}$ we set $H^1_{\mathcal{S}}(K_v, T) = H^1_{\mathrm{ur}}(K_v, T)$ and define the *Selmer group* associated to $\mathcal{S}$ as

$$H^1_{\mathcal{S}}(K, T) := \ker\left(H^1(K, T) \to \bigoplus_{v \in M_K} H^1(K_v, T)/H^1_{\mathcal{S}}(K_v, T) \right).$$

The following theorem, which is a very slight generalization of [loc. cit., Theorem 3.9], allows us to compare the dimensions of two Selmer groups modulo 2.

Theorem 6.6. *Let $\mathcal{S}$ and $\mathcal{S}'$ be two Selmer structures for $(T, \boldsymbol{q})$. Then*

$$\dim_{\mathbb{F}_p} H^1_{\mathcal{S}}(K, T) - \dim_{\mathbb{F}_p} H^1_{\mathcal{S}'}(K, T) \equiv \sum_{\Sigma_{\mathcal{S}} \cup \Sigma_{\mathcal{S}'}} \dim_{\mathbb{F}_p} H^1_{\mathcal{S}}(K_v, T)/(H^1_{\mathcal{S}}(K_v, T) \cap H^1_{\mathcal{S}'}(K_v, T)) \pmod 2.$$

Proof. This is proven for $\dim_{\mathbb{F}_p} T = 2$ in [loc. cit., Theorem 3.9] and the proof generalizes verbatim to the case where T has arbitrary (even) dimension with one subtlety: their proof relies on [loc. cit., Proposition 2.4] which is a general result concerning the dimension of the intersection of Lagrangian

subspaces of a finite dimensional metabolic space. The one difference from the case there is that now our quadratic forms (in general) take values in $\mathbb{Q}/\mathbb{Z}$ rather than just $\mathbb{F}_p$ as they assume. However, one readily verifies that this assumption is not used in the proof of the cited result. Alternatively, see [Česnavičius 2018, Theorem 5.9] which gives a further generalization of [Klagsbrun et al. 2013, Theorem 3.9] which includes our case. □

6F. ***Twisting data and twisted Selmer groups.*** Fix from now on a global metabolic structure $\boldsymbol{q}$ on T.

Definition 6.7. For each place $v \in M_K$, write $\mathcal{H}(q_v)$ for the set of Lagrangian subspaces for q_v and, for $v \notin \Sigma$, write $\mathcal{H}_{\mathrm{ram}}(q_v)$ for the subset of $\mathcal{H}(q_v)$ consisting of Lagrangian subspaces X for which $X \cap H^1_{\mathrm{ur}}(K_v, T) = 0$.

Definition 6.8 (twisting data). We define *twisting data* $\boldsymbol{\alpha}$ for $(T, \boldsymbol{q}, \Sigma)$ to consist of

(i) for each $v \in \Sigma$ a map

$$\alpha_v : \mathcal{F}(K_v) \to \mathcal{H}(q_v),$$

(ii) for each $v \notin \Sigma$ for which $\boldsymbol{\mu}_p \subseteq K_v$, a map

$$\alpha_v : \mathcal{F}_{\mathrm{ram}}(K_v) \to \mathcal{H}_{\mathrm{ram}}(q_v).$$

Remark 6.9. Our definition of twisting data is slightly different to that of [Klagsbrun et al. 2013, Definition 4.4]. In their case, since T has dimension 2, for $v \notin \Sigma$ and with $\boldsymbol{\mu}_p \subseteq K_v$, $\mathcal{H}_{\mathrm{ram}}(q_v)$ has cardinality 0,1, or p according to $\dim T^{G_{K_v}} = 0$, 1 or 2 respectively. In the first two cases they do not specify a map α_v as there is a unique such. In the final case they additionally insist that α_v is a bijection, as is possible since $\mathcal{F}_{\mathrm{ram}}(K_v)$ has order p.

Since for us T is allowed to have dimension greater that 2 we in general have $|\mathcal{H}_{\mathrm{ram}}(q_v)| > p$ and thus cannot insist that α_v is a bijection once it ceases to be unique. Although omitting this condition does not impact what follows, and is in fact not used in the main results of [Klagsbrun et al. 2013], we remark that it is used crucially in a follow up paper to that paper: [Klagsbrun et al. 2014].

Definition 6.10 (twisted Selmer groups). Let $(T, \boldsymbol{q}, \Sigma, \boldsymbol{\alpha})$ as above be fixed, and let $\chi \in \mathcal{C}(K)$. Let P_χ denote the set of primes of K for which χ ramifies. Then we define a Selmer structure $\mathcal{S}(\chi)$ by taking $\Sigma_{\mathcal{S}(\chi)}$ to be $\Sigma \cup P_\chi$ and setting $H^1_{\mathcal{S}(\chi)}(K_v, T) := \alpha_v(\chi_v)$ for $v \in \Sigma \cup P_\chi$. We write $\mathrm{Sel}(T, \chi)$ for the associated Selmer group

$$\mathrm{Sel}(T, \chi) := H^1_{\mathcal{S}(\chi)}(K, T).$$

6G. ***Comparing the parity of dimensions of twisted Selmer groups.*** From now on we fix T, the set of places Σ, a global metabolic structure $\boldsymbol{q}$, and twisting data $\boldsymbol{\alpha}$.

The following theorem, which is a slight variant of [Klagsbrun et al. 2013, Theorem 4.11] allows us to compare the parity of the dimensions of the Selmer groups $\mathrm{Sel}(T, \chi)$ as we vary χ. We first make one further definition.

Definition 6.11. Let v be a place of K and χ_1 and χ_2 be elements of $\mathcal{C}(K_v)$. Then we set

$$h_v(\chi_1, \chi_2) := \dim_{\mathbb{F}_p}(\alpha_v(\chi_1)/(\alpha_v(\chi_1) \cap \alpha_v(\chi_2))).$$

Note that since any two Lagrangian subspaces of $H^1(K_v, T)$ have the same dimension this is symmetric in χ_1 and χ_2.

Theorem 6.12. *For any $\chi \in \mathcal{C}(K)$ we have*

$$\dim_{\mathbb{F}_p} \mathrm{Sel}(T, \chi) - \dim_{\mathbb{F}_p} \mathrm{Sel}(T, \mathbb{1}_K) \equiv \sum_{v \in \Sigma} h_v(\mathbb{1}_{K_v}, \chi_v) + \sum_{v \notin \Sigma, \chi_v \text{ ram}} \dim_{\mathbb{F}_p} T^{G_{K_v}} \pmod 2$$

(*here the second sum is taken over places $v \notin \Sigma$ for which the character χ_v is ramified*).

Proof. This is essentially [Klagsbrun et al. 2013, Theorem 4.11]. Let $\mathcal{S}(\chi)$ and $\mathcal{S}(\mathbb{1}_K)$ be the Selmer structures associated to the characters χ and $\mathbb{1}_K$ respectively. Then

$$\Sigma_{\mathcal{S}(\chi)} \cup \Sigma_{\mathcal{S}(\mathbb{1}_K)} = \Sigma \sqcup \{v \notin \Sigma : \chi_v \text{ ramified}\}.$$

Applying Theorem 6.6 to $\mathcal{S}(\chi)$ and $\mathcal{S}(\mathbb{1}_K)$ and noting that, by the definition of the twisting data, $H^1_{\mathrm{ur}}(K_v, T) \cap \alpha_v(\chi_v) = 0$ for all $v \notin \Sigma$ for which χ_v is ramified, we obtain

$$\dim_{\mathbb{F}_p} \mathrm{Sel}(T, \chi) - \dim_{\mathbb{F}_p} \mathrm{Sel}(T, \mathbb{1}_K) \equiv \sum_{v \in \Sigma} h_v(\mathbb{1}_{K_v}, \chi_v) + \sum_{v \notin \Sigma, \chi_v \text{ ram}} \dim_{\mathbb{F}_p} H^1_{\mathrm{ur}}(K_v, T) \pmod 2.$$

The result now follows since for each $v \notin \Sigma$ we have $\dim_{\mathbb{F}_p} H^1_{\mathrm{ur}}(K_v, T) = \dim_{\mathbb{F}_p} T^{G_{K_v}}$. This is shown in (the proof of) [Klagsbrun et al. 2013, Lemma 3.7] in the case that T has dimension 2. The general case is identical. □

7. Disparity in Selmer ranks: statement and first cases

In this section we fix $(T, \Sigma, \boldsymbol{q}, \boldsymbol{\alpha})$ as in the previous section and consider the proportion of characters χ for which the associated Selmer groups $\mathrm{Sel}(T, \chi)$ have odd (resp. even) $\mathbb{F}_p$-dimension. To make this precise, one has to order the elements of $\mathcal{C}(K)$.

7A. *Ordering twists.*

We use the same ordering as in [Klagsbrun et al. 2013, Definition 7.3].

Definition 7.1. For $\chi \in \mathcal{C}(K)$, set

$$\|\chi\| = \max\{N(\mathfrak{p}) : \chi \text{ is ramified at } \mathfrak{p}\}$$

(where here for a prime $\mathfrak{p} \lhd \mathcal{O}_K$, $N(\mathfrak{p})$ denotes the norm of $\mathfrak{p}$). If this set is empty, our convention is that $\|\chi\| = 1$. Now for each $X > 0$ define

$$\mathcal{C}(K, X) = \{\chi \in \mathcal{C}(K) : \|\chi\| < X\}.$$

For each $X \geq 1$ this is a finite subgroup of $\mathcal{C}(K)$ and each element of $\mathcal{C}(K)$ appears in $\mathcal{C}(K, X)$ for some X. We will make crucial use of the group structure on the $\mathcal{C}(K, X)$ to facilitate with counting problems.

We will repeatedly use the following fact.

Lemma 7.2. *For all sufficiently large $X > 0$ the restriction homomorphism*

$$\mathcal{C}(K, X) \to \prod_{v \in \Sigma} \mathcal{C}(K_v)$$

sending χ to $(\chi_v)_{v \in \Sigma}$ is surjective.

Proof. This follows immediately from the Grunwald–Wang theorem. See for example [Neukirch et al. 2008, Theorem 9.2.3(ii)]. See also [Klagsbrun et al. 2013, Proposition 6.8(i)] but note that they have a running hypothesis on the set of places Σ which we do not wish to impose at this stage. □

7B. ***Statement of the result.*** The proportion of characters for which $\dim_{\mathbb{F}_p} \mathrm{Sel}(T, \chi)$ is even (resp. odd) will depend heavily on the action of G_K on T. More specifically, it will depend on the behavior of the following function. Recall that $K(T)$ denotes the field of definition of the elements of T.

Definition 7.3. Write $G := \mathrm{Gal}(K(T)/K)$ and define the function:

$$\begin{aligned} \epsilon : G &\to \{\pm 1\} \\ \sigma &\mapsto (-1)^{\dim_{\mathbb{F}_p} T^\sigma} \end{aligned}$$

The result is then the following.

Theorem 7.4. *We have*:

(i) *If either $p = 2$ and ϵ fails to be a homomorphism, or $p > 2$ and ϵ is nontrivial when restricted to* $\mathrm{Gal}(K(T)/K(\boldsymbol{\mu}_p))$, *then*

$$\lim_{X \to \infty} \frac{|\{\chi \in \mathcal{C}(K, X) : \dim_{\mathbb{F}_p} \mathrm{Sel}(T, \chi) \text{ is even}\}|}{|\mathcal{C}(K, X)|} = \frac{1}{2}.$$

Moreover, if $p = 2$ then it suffices to take X sufficiently large as opposed to taking the limit $X \to \infty$.

(ii) *If either $p = 2$ and ϵ is a homomorphism, or $p > 2$ and ϵ is trivial when restricted to* $\mathrm{Gal}(K(T)/K(\boldsymbol{\mu}_p))$, *then for all sufficiently large X we have*

$$\frac{|\{\chi \in \mathcal{C}(K, X) : \dim_{\mathbb{F}_p} \mathrm{Sel}(T, \chi) \text{ is even}\}|}{|\mathcal{C}(K, X)|} = \frac{1 + (-1)^{\dim_{\mathbb{F}_p} \mathrm{Sel}(T, \mathbb{1}_K)} \cdot \delta}{2}$$

with $\delta = \prod_{v \in \Sigma} \delta_v$ given in Definition 7.8.

The proof of Theorem 7.4, which is a combination of Theorems 7.10 and 9.5, will occupy the remainder of Sections 7–9.

Remark 7.5. Here we briefly discuss the function ϵ. For convenience we identify $\boldsymbol{\mu}_p$ with the additive group of $\mathbb{F}_p$ and think of the pairing $(\cdot,\cdot)$ as landing in $\mathbb{F}_p$. Due to this pairing, the group $G = \mathrm{Gal}(K(T)/K)$ is a subgroup of the general symplectic group

$$\mathrm{GSp}(T) = \{g \in \mathrm{GL}(V) : \forall v, w \in T,\ (gv, gw) = \lambda(g)(v, w) \text{ for some } \lambda(g) \in \mathbb{F}_p^\times\}.$$

First suppose $p = 2$ so that $\mathrm{GSp}(T) = \mathrm{Sp}(T)$ is the symplectic group associated to $(\cdot,\cdot)$. If $\dim_{\mathbb{F}_2} T > 4$ then $\mathrm{Sp}(T)$ is simple and since any symplectic transvection σ (i.e., element of $\mathrm{Sp}(T)$ of the form $v \mapsto v + (v, w)w$ for fixed $0 \neq w \in T$) has $\dim_{\mathbb{F}_2} T^\sigma$ odd, if G is isomorphic to $\mathrm{Sp}(T)$ (i.e., is as large as possible) then ϵ is not a homomorphism. Thus case (i) of Theorem 7.4 is, in some sense, the "generic" case. When $\dim_{\mathbb{F}_2} T = 2$ one can check that ϵ is always a homomorphism, whilst if $\dim_{\mathbb{F}_2} T = 4$ then $\mathrm{Sp}(T)$ is isomorphic to the symmetric group S_6. One can check (see Example 10.17 later) that when G is either the whole of S_6 or the alternating group A_6 then ϵ is not a homomorphism, so again case (i) of Theorem 7.4 holds for G "large enough". On the other hand, Proposition 3.5 gives a supply of examples where ϵ is a homomorphism. Namely, if G fixes a quadratic refinement q of $(\cdot,\cdot)$ then G is a subgroup of the orthogonal group $O(q)$, in which case ϵ is the Dickson homomorphism.

Now suppose that $p > 2$. The subgroup $\mathrm{Gal}(K(T)/K(\boldsymbol{\mu}_p))$ consists of those elements $g \in G$ for which $\lambda(g) = 1$. That is, it is the intersection of G with the symplectic group $\mathrm{Sp}(T)$. If G contains a symplectic transvection σ (which as now $p > 2$ is an element of $\mathrm{Sp}(T)$ of the form $v \mapsto v + \beta \cdot (v, w)w$ for $\beta \in \mathbb{F}_p^\times$, $0 \neq w \in T$) then one sees easily that $\epsilon(\sigma) = -1$, so that ϵ is nontrivial when restricted to $\mathrm{Gal}(K(T)/K(\boldsymbol{\mu}_p))$. Thus again case (i) of Theorem 7.4 holds for G "large enough".

7C. ***The cases $p = 2$ and ϵ is a homomorphism, and $p > 2$ and ϵ is trivial when restricted to $\mathrm{Gal}(K(T)/K(\boldsymbol{\mu}_p))$.*** Suppose now that either $p = 2$ and ϵ is a homomorphism, or $p > 2$ and ϵ is trivial for all $\sigma \in \mathrm{Gal}(K(T)/K(\boldsymbol{\mu}_p))$.

Definition 7.6. Let $v \in \Sigma$ and $\chi \in \mathcal{C}(K_v)$. If $p > 2$ we define

$$\omega_v(\chi) := (-1)^{h_v(\mathbb{1}_{K_v}, \chi)}.$$

If $p = 2$ view ϵ as a quadratic character of K and let $\Delta \in K^\times/K^{\times 2}$ be such that the corresponding quadratic extension is given by $K(\sqrt{\Delta})/K$. We then define

$$\omega_v(\chi) := \chi(\Delta)(-1)^{h_v(\mathbb{1}_{K_v}, \chi)}$$

where here for a place v of K we evaluate χ_v at Δ via local class field theory.

Lemma 7.7. *For any $\chi \in \mathcal{C}(K)$ we have*

$$(-1)^{\dim_{\mathbb{F}_p} \mathrm{Sel}(T,\chi)} = (-1)^{\dim_{\mathbb{F}_p} \mathrm{Sel}(T,\mathbb{1}_K)} \prod_{v \in \Sigma} \omega_v(\chi_v).$$

Proof. Fix $v \notin \Sigma$ with $\boldsymbol{\mu}_p \subseteq K_v$, and let $\mathrm{Frob}_v \in G$ denote the Frobenius element at v in $K(T)/K$. Then as T is unramified at v we have

$$(-1)^{\dim_{\mathbb{F}_p} T^{G_{K_v}}} = \epsilon(\mathrm{Frob}_v).$$

If $p > 2$ then $\epsilon(\mathrm{Frob}_v) = 1$ for all $v \notin \Sigma$ by assumption, whence the result follows from Theorem 6.12.

Now suppose that $p = 2$. As above, view ϵ as a quadratic character of K. Since ϵ factors through $\mathrm{Gal}(K(T)/K)$ it is unramified outside Σ. In particular, if $v \notin \Sigma$ is such that χ_v is unramified, then both ϵ_v and χ_v are unramified at v and so $\chi_v(\Delta) = 1$. On the other hand, if $v \notin \Sigma$ is such that χ_v ramifies at v then since K_v has odd residue characteristic, we have $\chi_v(\Delta) = \epsilon(\mathrm{Frob}_v)$ (see Lemma 8.4(ii)). Global class field theory gives $\prod_{v \in M_K} \chi_v(\Delta) = 1$ from which it follows that

$$\prod_{v \notin \Sigma,\, \chi_v \text{ ram}} (-1)^{\dim_{\mathbb{F}_p} T^{G_{K_v}}} = \prod_{v \in \Sigma} \chi_v(\Delta).$$

We now conclude by Theorem 6.12. □

The proof of Theorem 7.4(ii) now proceeds as in [Klagsbrun et al. 2013, §7].

Definition 7.8. For each $v \in \Sigma$ define

$$\delta_v := \frac{1}{|\mathcal{C}(K_v)|} \sum_{\chi \in \mathcal{C}(K_v)} \omega_v(\chi) \quad \text{and} \quad \delta := \prod_{v \in \Sigma} \delta_v.$$

Remark 7.9. We have decided to define δ slightly differently to [Klagsbrun et al. 2013, §7] so that it is a product of local terms. Our definition of the δ_v is consistent with theirs however.

Theorem 7.10. *Suppose that either $p = 2$ and ϵ is a homomorphism, or $p > 2$ and ϵ is trivial when restricted to* $\mathrm{Gal}(K(T)/K(\boldsymbol{\mu}_p))$. *Then for all sufficiently large $X > 0$ we have*

$$\frac{|\{\chi \in \mathcal{C}(K, X) : \dim_{\mathbb{F}_p} \mathrm{Sel}(T, \chi) \text{ is even}\}|}{|\mathcal{C}(K, X)|} = \frac{1 + (-1)^{\dim_{\mathbb{F}_p} \mathrm{Sel}(T, \mathbb{1}_K)} \delta}{2}.$$

Proof. The argument is the same as in [Klagsbrun et al. 2013, Theorem 7.6]. We repeat it for convenience. Write $\Gamma = \prod_{v \in \Sigma} \mathcal{C}(K_v)$ and for $\chi \in \mathcal{C}(K)$, write $\chi|_\Gamma$ for the image of χ under the natural restriction homomorphism $\mathcal{C}(K) \to \Gamma$ sending χ to $(\chi_v)_{v \in \Sigma}$. From Lemma 7.7 we see that the parity of $\dim_{\mathbb{F}_p} \mathrm{Sel}(T, \chi)$ depends only on $\chi|_\Gamma$ and that $\dim_{\mathbb{F}_p} \mathrm{Sel}(T, \chi)$ is even if and only if

$$\prod_{v \in \Sigma} \omega(\chi_v) = (-1)^{\dim_{\mathbb{F}_p} \mathrm{Sel}(T, \mathbb{1}_K)}.$$

As is possible by Lemma 7.2, take X sufficiently large that $\mathcal{C}(K, X)$ surjects onto Γ under restriction. Since restriction is a group homomorphism, its fibers all have the same size (being cosets of the kernel) and, in particular, we have

$$\frac{|\{\chi \in \mathcal{C}(K, X) : \dim_{\mathbb{F}_p} \mathrm{Sel}(T, \chi) \text{ is even}\}|}{|\mathcal{C}(K, X)|} = \frac{|\{\gamma \in \Gamma : \prod_{v \in \Sigma} \omega(\gamma_v) = (-1)^{\dim_{\mathbb{F}_p} \mathrm{Sel}(T, \mathbb{1}_K)}\}|}{|\Gamma|}$$

where here, for $\gamma \in \Gamma$ we denote by γ_v its projection onto $\mathcal{C}(K_v)$.

To evaluate the right-hand side of the above expression, define

$$N := \left| \left\{ \gamma \in \Gamma : \prod_{v \in \Sigma} \omega(\gamma_v) = 1 \right\} \right|.$$

Then we have

$$N - (|\Gamma| - N) = \sum_{\gamma\in\Gamma}\prod_{v\in\Sigma}\omega(\gamma_v) = \prod_{v\in\Sigma}\sum_{\chi_v\in\mathcal{C}(K_v)}\omega(\chi_v).$$

Dividing the above expression through by $2|\Gamma|$ gives

$$\frac{|\{\gamma\in\Gamma : \prod_{v\in\Sigma}\omega(\gamma_v) = 1\}|}{|\Gamma|} = \frac{1+\delta}{2}$$

and the result follows immediately. □

8. Disparity in Selmer ranks: local symbols and global characters

In order to prove the remaining cases of Theorem 7.4 we now recall and slightly generalize (as well as rephrase for convenience in Section 9) the results of [Klagsbrun et al. 2013, §6], which uses class field theory to analyze which collections of local characters arise from a global character.

8A. ***Local symbols.*** For each nonarchimedean place v of K, Tate local duality gives a nondegenerate pairing

$$H^1(K_v, \boldsymbol{\mu}_p) \times H^1(K_v, \mathbb{Z}/p\mathbb{Z}) \to \mathbb{Q}/\mathbb{Z}, \tag{8.1}$$

defined as the composition

$$H^1(K_v, \boldsymbol{\mu}_p) \times H^1(K_v, \mathbb{Z}/p\mathbb{Z}) \xrightarrow{\cup} H^2(K, \boldsymbol{\mu}_p) \hookrightarrow \mathrm{Br}(K_v) \xrightarrow{\mathrm{inv}_v} \mathbb{Q}/\mathbb{Z}$$

(here the map "$\cup$" is the cup product map on cohomology combined with the canonical isomorphism $\mathbb{Z}/p\mathbb{Z}\otimes\boldsymbol{\mu}_p \cong \boldsymbol{\mu}_p$).

We now slightly modify this pairing. As the Galois action on $\mathbb{Z}/p\mathbb{Z}$ is trivial we have $H^1(K_v, \mathbb{Z}/p\mathbb{Z}) = \mathrm{Hom}_{\mathrm{cnt}}(G_{K_v}, \mathbb{Z}/p\mathbb{Z})$. Picking an isomorphism of abstract groups $\theta : \boldsymbol{\mu}_p \xrightarrow{\sim} \mathbb{Z}/p\mathbb{Z}$ induces isomorphisms

$$\mathcal{C}(K_v) \cong H^1(K_v, \mathbb{Z}/p\mathbb{Z}) \quad \text{and} \quad \tfrac{1}{p}\mathbb{Z}/\mathbb{Z} \cong \boldsymbol{\mu}_p \tag{8.2}$$

where for the latter we identify $\mathbb{Z}/p\mathbb{Z}$ with $\frac{1}{p}\mathbb{Z}/\mathbb{Z}$ by sending $1\in\mathbb{Z}/p\mathbb{Z}$ to $\frac{1}{p}$. Noting that $H^2(K_v, \boldsymbol{\mu}_p)\subseteq \mathrm{Br}(K_v)$ is mapped by inv_v into $\frac{1}{p}\mathbb{Z}/\mathbb{Z}$, combining the pairing (8.1) with the isomorphisms of (8.2) yields a nondegenerate pairing

$$[\,\cdot\,,\cdot\,]_v : H^1(K_v, \boldsymbol{\mu}_p) \times \mathcal{C}(K_v) \to \boldsymbol{\mu}_p \tag{8.3}$$

which is easily seen to be independent of the choice of θ.

The following well-known lemma summarizes the properties of this local pairing.

Lemma 8.4. *Let v be a nonarchimedean place of K. Then*:

(i) *If $v\nmid p$ then the groups $H^1_{\mathrm{ur}}(K_v, \boldsymbol{\mu}_p)$ and $\mathcal{C}_{\mathrm{ur}}(K_v)$ are orthogonal complements with respect to the pairing $[\,\cdot\,,\cdot\,]_v$.*

(ii) *Let $x \in K_v^\times$ and write $\phi_x \in H^1(K_v, \boldsymbol{\mu}_p)$ for the image of x under the boundary map associated to the Kummer sequence*

$$1 \to \boldsymbol{\mu}_p \to \overline{K_v}^\times \xrightarrow{x \mapsto x^p} \overline{K_v}^\times \to 1.$$

Then for any $\chi \in \mathcal{C}(K_v)$ we have

$$[\phi_x, \chi]_v = \chi(\mathrm{Art}_{K_v}(x))^{-1},$$

where here $\mathrm{Art}_{K_v} : K_v^\times \to G_{K_v}^{\mathrm{ab}}$ denotes the local Artin map.

(iii) *Suppose v is such that $\boldsymbol{\mu}_p \subseteq K_v$ so that $H^1(K_v, \boldsymbol{\mu}_p) = \mathcal{C}(K_v)$. Then the resulting pairing*

$$[\,\cdot\,,\cdot\,]_v : \mathcal{C}(K_v) \times \mathcal{C}(K_v) \to \boldsymbol{\mu}_p$$

is antisymmetric.

Proof. Part (i) is [Neukirch et al. 2008, Theorem 7.2.15] whilst part (ii) is Corollary 7.2.13 of [loc. cit.]. (The cited results are stated for the pairing of (8.1) rather than the altered pairing $[\,\cdot\,,\cdot\,]_v$ but in each case they immediately imply the claimed results.) Finally, antisymmetry of the cup product

$$H^1(K_v, \mathbb{Z}/p\mathbb{Z}) \times H^1(K_v, \mathbb{Z}/p\mathbb{Z}) \to H^2(K_v, \mathbb{Z}/p\mathbb{Z} \otimes \mathbb{Z}/p\mathbb{Z})$$

gives part (iii). □

8B. ***Existence of global characters with specified restriction and ramification.*** We will need the following lemma which is the analogue of [Klagsbrun et al. 2013, Proposition 6.8(iii)] in the case that the dimension of T is allowed to be larger than 2.

Notation 8.5. Writing $\Gamma := \prod_{v \in \Sigma} \mathcal{C}(K_v)$, we denote by $[\,\cdot\,,\cdot\,]_\Sigma$ the nondegenerate bilinear pairing

$$[\,\cdot\,,\cdot\,]_\Sigma : \left(\prod_{v \in \Sigma} H^1(K_v, \boldsymbol{\mu}_p) \right) \times \Gamma \to \boldsymbol{\mu}_p$$

defined as the sum (or rather product) over $v \in \Sigma$ of the pairings $[\,\cdot\,,\cdot\,]_v$ of (8.3).

Lemma 8.6. *Let P denote the set of primes of K not in Σ which split completely in $K(T)/K$, and fix $\gamma \in \Gamma$. Then there is a character $\chi \in \mathcal{C}(K)$ unramified outside $\Sigma \cup P$ and with $\chi|_\Gamma = \gamma$, if and only if $[c, \gamma]_\Sigma = 0$ for each c in the image of the restriction homomorphism*

$$H^1(K(T)/K, \boldsymbol{\mu}_p) \to \prod_{v \in \Sigma} H^1(K_v, \boldsymbol{\mu}_p).$$

Proof. Exactness at the middle term of the Poitou–Tate exact sequence (see, for example, [Milne 2006, Theorem I.4.10]) applied to the set $\Sigma \cup P$ of places and the G_K-module $\mathbb{Z}/p\mathbb{Z}$ (and its dual $\boldsymbol{\mu}_p$), shows that

$$\mathrm{im}\left(H^1(K_{\Sigma \cup P}/K, \mathbb{Z}/p\mathbb{Z}) \to \prod_{v \in \Sigma \cup P}{}' H^1(K_v, \mathbb{Z}/p\mathbb{Z}) \right)$$

is the orthogonal complement of

$$\operatorname{im}\left(H^1(K_{\Sigma\cup P}/K, \boldsymbol{\mu}_p) \to \prod_{v\in\Sigma\cup P}{}' H^1(K_v, \boldsymbol{\mu}_p)\right)$$

under the sum of the local pairings of (8.1), where here $K_{\Sigma\cup P}$ denotes the maximal extension of K unramified outside $\Sigma\cup P$ and the restricted direct products are taken with respect to unramified classes.

Now fix any choice of isomorphism $\boldsymbol{\mu}_p \cong \mathbb{Z}/p\mathbb{Z}$ and use it to identify $\mathcal{C}(K)$ with $H^1(K, \mathbb{Z}/p\mathbb{Z})$, and $\mathcal{C}(K_v)$ with $H^1(K_v, \mathbb{Z}/p\mathbb{Z})$ for each v similarly. Then the group $H^1(K_{\Sigma\cup P}/K, \mathbb{Z}/p\mathbb{Z})$ corresponds to the group of characters unramified outside $\Sigma\cup P$, which we denote by $\mathcal{C}(K)_{\Sigma\cup P}$. Making these identifications and projecting onto $\prod_{v\in\Sigma}\mathcal{C}(K_v)$, it follows formally that the image of $\mathcal{C}(K)_{\Sigma\cup P}$ in $\prod_{v\in\Sigma}\mathcal{C}(K_v)$ is the orthogonal complement with respect to the pairing $[\,\cdot\,,\,\cdot\,]_\Sigma$ of the image of

$$\ker\left(H^1(K_{\Sigma\cup P}/K, \boldsymbol{\mu}_p) \to \prod_{v\in P}{}' H^1(K_v, \boldsymbol{\mu}_p)\right)$$

in $\prod_{v\in\Sigma} H^1(K_v, \boldsymbol{\mu}_p)$. We now conclude by the following lemma. □

Lemma 8.7. *Let P denote the set of primes of K not in Σ and which split completely in $K(T)/K$, and let $K_{\Sigma\cup P}$ denote the maximal extension of K unramified outside $\Sigma\cup P$. Then we have*

$$H^1(K(T)/K, \boldsymbol{\mu}_p) = \ker\left(H^1(K_{\Sigma\cup P}/K, \boldsymbol{\mu}_p) \to \prod_{v\in P}{}' H^1(K_v, \boldsymbol{\mu}_p)\right),$$

the groups being compared inside $H^1(K, \boldsymbol{\mu}_p)$ (and the restricted direct product being taken with respect to unramified classes as above).

Proof. Since $K(T)$ is unramified outside Σ we have $K(T)\subseteq K_{\Sigma\cup P}$. Thus it suffices to show that we have

$$H^1(K(T)/K, \boldsymbol{\mu}_p) = \ker\left(H^1(K, \boldsymbol{\mu}_p) \xrightarrow{\text{res}} \prod_{v\in P}{}' H^1(K_v, \boldsymbol{\mu}_p)\right).$$

Since each prime in P splits completely in $K(T)/K$ the restriction map above factors as

$$H^1(K, \boldsymbol{\mu}_p) \xrightarrow{f_1} H^1(K(T), \boldsymbol{\mu}_p) \xrightarrow{f_2} \prod_{v\in\mathcal{P}}{}' H^1(K_v, \boldsymbol{\mu}_p),$$

where both maps are given by restriction. Since the inflation-restriction exact sequence identifies $H^1(K(T)/K, \boldsymbol{\mu}_p)$ with $\ker(f_1)$, it suffices to show that f_2 is injective. Since $K(T)$ and each K_v ($v\in P$) contain $\boldsymbol{\mu}_p$, we may reinterpret f_2 as the restriction map on characters

$$\mathcal{C}(K(T)) \to \prod_{v\in P}{}' \mathcal{C}(K_v).$$

Suppose $\chi\in\mathcal{C}(K(T))$ is a character of $K(T)$ which is trivial in $\mathcal{C}(K_v)$ for each $v\in P$, let $L/K(T)$ denote the extension corresponding to the fixed field of the kernel of χ, and let L'/K denote the Galois closure of L/K. Then our assumption on χ means that every prime $v\in\mathcal{P}$ splits completely in L'/K. By

the Chebotarev density theorem this gives $[L' : K] \leq [K(T) : K]$. Since we already know that $K(T) \subseteq L'$ we must have $L' = K(T)$ whence χ is the trivial character. □

8C. ***Assumptions on the set of places*** Σ**.** We now impose conditions on the finite set of places Σ (in addition to containing all archimedean places, all primes over p and all places for which T is ramified) which will be necessary for the proof of the remaining cases of Theorem 7.4.

Assumption 8.8. We henceforth impose the following conditions on the finite set of places Σ:

(i) The restriction homomorphism

$$H^1(K(T)/K, \boldsymbol{\mu}_p) \to \prod_{v \in \Sigma} H^1(K_v, \boldsymbol{\mu}_p)$$

is injective.

(ii) $\mathrm{Pic}(\mathcal{O}_{K,\Sigma}) = 0$.

(iii) The natural map

$$\mathcal{O}_{K,\Sigma}^\times / (\mathcal{O}_{K,\Sigma}^\times)^p \to \prod_{v \in \Sigma} K_v^\times / (K_v^\times)^p$$

is injective.

(In (ii) and (iii), $\mathcal{O}_{K,\Sigma}$ denotes the elements of K integral outside Σ.)

Lemma 8.9. *A set of places Σ satisfying Assumption 8.8 exists.*

Proof. We begin by taking Σ large enough that it contains all archimedean places, all primes over p and all places where T ramifies. By the Grunwald–Wang theorem [Neukirch et al. 2008, Theorem 9.1.9(ii)] the map

$$H^1(K, \boldsymbol{\mu}_p) \to \prod_{v \in M_K} H^1(K_v, \boldsymbol{\mu}_p)$$

is injective. In particular, as $H^1(K(T)/K, \boldsymbol{\mu}_p)$ is a finite subgroup of $H^1(K, \boldsymbol{\mu}_p)$, we see that by enlarging Σ if necessary we may additionally ensure that (i) holds.

Finally, [Klagsbrun et al. 2013, Lemma 6.1] shows that any finite set of places may be further enlarged so that (ii) and (iii) hold. □

Lemma 8.10. *Suppose Assumption 8.8 is satisfied and let $\mathfrak{p}$ be a prime of K with $\mathfrak{p} \notin \Sigma$ and $\boldsymbol{\mu}_p \subseteq K_\mathfrak{p}^\times$. Write*

$$\delta_\mathfrak{p} : K_\mathfrak{p}^\times / K_\mathfrak{p}^{\times p} \xrightarrow{\sim} \mathcal{C}(K_\mathfrak{p})$$

for the isomorphism (coming from the Kummer sequence) sending $x \in K_\mathfrak{p}^\times$ to the character $\sigma \mapsto \sigma(y)/y$ where $y \in \overline{K}_\mathfrak{p}^\times$ is such that $y^p = x$ (any two choices for y yield the same character since $\boldsymbol{\mu}_p \subseteq K_\mathfrak{p}$).

Then there is a (global) character $\varphi(\mathfrak{p}) \in \mathcal{C}(K)$ satisfying the following three conditions:

$$\begin{aligned} &\varphi(\mathfrak{p}) \text{ ramifies at } \mathfrak{p}. \\ &\varphi(\mathfrak{p}) \text{ is unramified outside } \Sigma \cup \{\mathfrak{p}\}. \\ &\text{The restriction of } \varphi(\mathfrak{p}) \text{ to } \mathcal{C}(K_\mathfrak{p}) \text{ is equal to } \delta_\mathfrak{p}(\varpi) \text{ for some uniformizer } \varpi \text{ of } K_\mathfrak{p}. \end{aligned} \tag{8.11}$$

Proof. Given the assumptions on Σ, the existence of a character $\varphi(\mathfrak{p})$ which ramifies at $\mathfrak{p}$ and is unramified outside $\Sigma \cup \{\mathfrak{p}\}$ follows from [Klagsbrun et al. 2013, Proposition 6.8 (ii)]. Fix one such and pick $x \in K_{\mathfrak{p}}^{\times}$ such that the restriction of $\varphi(\mathfrak{p})$ is equal to $\delta_{\mathfrak{p}}(x)$. Since $\varphi(\mathfrak{p})$ ramifies at $\mathfrak{p}$ the extension of $K_{\mathfrak{p}}^{\times}$ obtained by adjoining a p-th root of x ramifies. In particular, since $K_{\mathfrak{p}}$ has residue characteristic coprime to p (as $\mathfrak{p} \notin \Sigma$), the valuation $v_{\mathfrak{p}}(x)$ of x is coprime to p. Noting that replacing $\varphi(\mathfrak{p})$ with $\varphi(\mathfrak{p})^m$ for any m coprime to p yields another character which ramifies at $\mathfrak{p}$ and is unramified outside $\Sigma \cup \{\mathfrak{p}\}$, we may suppose that $v_{\mathfrak{p}}(x)$ is congruent to 1 modulo p. Finally, since $K_{\mathfrak{p}}^{\times p}$ is in the kernel of $\delta_{\mathfrak{p}}$ we may now shift x by a p-th power of a uniformizer to suppose that x has valuation 1 as desired. □

The following lemma evaluates the pairing $[\,\cdot\,,\,\cdot\,]_{\Sigma}$ of Notation 8.5 between the characters $\varphi(\mathfrak{p})$ of Lemma 8.10 and elements of $H^1(K(T)/K, \boldsymbol{\mu}_p)$.

Lemma 8.12. *Let $\mathfrak{p}$ be a prime of K not in Σ, let $\varphi(\mathfrak{p})$ satisfy (8.11), and let $c \in H^1(K(T)/K, \boldsymbol{\mu}_p)$. Then writing $\mathrm{Frob}_{\mathfrak{p}}$ for the Frobenius element at $\mathfrak{p}$ in $\mathrm{Gal}(K(T)/K)$ we have*

$$[c, \varphi(\mathfrak{p})]_{\Sigma} = c(\mathrm{Frob}_{\mathfrak{p}}).$$

Proof. By global class field theory the product of $[c, \varphi(\mathfrak{p})]_v$ over all places of K is equal to 1. In particular, we have

$$[c, \varphi(\mathfrak{p})]_{\Sigma} = \prod_{v \notin \Sigma} [c, \varphi(\mathfrak{p})]_v.$$

If $\mathfrak{q}$ is a prime of K not in Σ then $\mathfrak{q} \nmid p$ and, additionally, $K(T)/K$ is unramified at $\mathfrak{q}$ whence the restriction of c to $H^1(K_{\mathfrak{q}}, \boldsymbol{\mu}_p)$ is in the unramified subgroup $H^1_{\mathrm{ur}}(K_{\mathfrak{q}}, \boldsymbol{\mu}_p)$. If $\mathfrak{q} \neq \mathfrak{p}$ then $\varphi(\mathfrak{p})$ is also unramified at $\mathfrak{q}$ whence $[c, \varphi(\mathfrak{p})]_{\mathfrak{q}} = 1$ by Lemma 8.4(i).

It now follows that $[c, \varphi(\mathfrak{p})]_{\Sigma} = [c, \varphi(\mathfrak{p})]_{\mathfrak{p}}$ and to conclude we must show that $[c, \varphi(\mathfrak{p})]_{\mathfrak{p}} = c(\mathrm{Frob}_{\mathfrak{p}})$. Since $\boldsymbol{\mu}_p \subseteq K_{\mathfrak{p}}$ and we've chosen $\varphi(\mathfrak{p})$ so that its restriction to $\mathcal{C}(K_{\mathfrak{p}})$ agrees with $\delta_{\mathfrak{p}}(\varpi)$ for some uniformizer ϖ of $K_{\mathfrak{p}}$, parts (ii) and (iii) of Lemma 8.4 combine to give

$$[c, \varphi(\mathfrak{p})]_{\mathfrak{p}} = c(\mathrm{Art}_{K_{\mathfrak{p}}}(\varpi)).$$

Now c is unramified at $\mathfrak{p}$ and by standard properties of the local Artin map we have $\mathrm{Art}_{K_{\mathfrak{p}}}(\varpi)|_{K_{\mathfrak{p}}^{\mathrm{nr}}} = \mathrm{Frob}_{K_{\mathfrak{p}}}$. On the other hand, since c came from $H^1(K(T)/K, \boldsymbol{\mu}_p)$, its restriction to $H^1(K_{\mathfrak{p}}, \boldsymbol{\mu}_p)$ factors through $\mathrm{Gal}(K_{\mathfrak{p}}(T)/K_{\mathfrak{p}})$. As the restriction of $\mathrm{Frob}_{K_{\mathfrak{p}}}$ to $\mathrm{Gal}(K_{\mathfrak{p}}(T)/K_{\mathfrak{p}})$ is precisely $\mathrm{Frob}_{\mathfrak{p}}$, we have the result. □

9. Disparity in Selmer ranks: remaining cases

We now treat the remaining cases of Theorem 7.4, namely when $p = 2$ and ϵ fails to be a homomorphism, or when $p > 2$ and ϵ is nontrivial when restricted to $\mathrm{Gal}(K(T)/K(\boldsymbol{\mu}_p))$. Our strategy is broadly based on that of [Klagsbrun et al. 2013, §8], although the arguments are more involved.

We begin by fixing a finite set of places Σ satisfying Assumption 8.8. As before let G denote the Galois group of $K(T)/K$ and write $\Gamma := \prod_{v \in \Sigma} \mathcal{C}(K_v)$. For $\chi \in \mathcal{C}(K)$ we denote by $\chi|_{\Gamma}$ the image of χ in Γ under the (product of the) natural restriction map(s).

Definition 9.1. Define a map $w : \mathcal{C}(K) \to \{\pm 1\}$ by

$$w(\chi) := \prod_{v \notin \Sigma, \chi_v \text{ ram}} (-1)^{\dim_{\mathbb{F}_p} T^{G_{K_v}}} = \prod_{v \notin \Sigma, \chi_v \text{ ram}} \epsilon(\mathrm{Frob}_v),$$

where here $\mathrm{Frob}_v \in G$ denotes the Frobenius element at v in $K(T)/K$.

Remark 9.2. By Theorem 6.12, for each $\chi \in \mathcal{C}(K)$ we have

$$(-1)^{\dim_{\mathbb{F}_2} \mathrm{Sel}(T,\chi)} = w(\chi)(-1)^{\dim_{\mathbb{F}_2} \mathrm{Sel}(T,\mathbb{1}_K)} \prod_{v \in \Sigma} (-1)^{h_v(\mathbb{1}_K, \chi_v)}.$$

We now examine the extent to which $w(\chi)$ behaves "independently" of the restriction of χ to Γ. To this end, we make the following definition.

Definition 9.3. For each $X \geq 1$ and $\gamma \in \Gamma$, define

$$s_X(\gamma) = \frac{|\{\chi \in \mathcal{C}(K, X) : \chi|_\Gamma = \gamma, w(\chi) = 1\}|}{|\{\chi \in \mathcal{C}(K, X) : \chi|_\Gamma = \gamma\}|}.$$

The rest of the section is occupied with the proof of the following theorem.

Theorem 9.4. *We have*:

(i) *If $p = 2$ and ϵ fails to be a homomorphism then, for all sufficiently large X, $s_X(\gamma) = \frac{1}{2}$ for all $\gamma \in \Gamma$.*

(ii) *If $p > 2$ and ϵ is nontrivial when restricted to $\mathrm{Gal}(K(T)/K(\boldsymbol{\mu}_p))$ then $\lim_{X\to\infty} s_X(\gamma) = \frac{1}{2}$ for all $\gamma \in \Gamma$.*

Assuming this for the moment we get as a corollary the remaining cases of Theorem 7.4.

Theorem 9.5. *We have*:

(i) *If $p = 2$ and ϵ fails to be a homomorphism then, for all sufficiently large X.*

$$\frac{|\{\chi \in \mathcal{C}(K, X) : \dim_{\mathbb{F}_p} \mathrm{Sel}(T, \chi) \text{ is even}\}|}{|\mathcal{C}(K, X)|} = \frac{1}{2},$$

(ii) *If $p > 2$ and ϵ is nontrivial when restricted to $\mathrm{Gal}(K(T)/K(\boldsymbol{\mu}_p))$ then*

$$\lim_{X\to\infty} \frac{|\{\chi \in \mathcal{C}(K, X) : \dim_{\mathbb{F}_p} \mathrm{Sel}(T, \chi) \text{ is even}\}|}{|\mathcal{C}(K, X)|} = \frac{1}{2}.$$

Proof. Fix $\gamma \in \Gamma$ and suppose that $\chi \in \mathcal{C}(K, X)$ is such that $\chi|_\Gamma = \gamma$. Then by Remark 9.2 we have

$$\dim_{\mathbb{F}_2} \mathrm{Sel}(T, \chi) \text{ is even} \Leftrightarrow w(\chi) = (-1)^{\dim_{\mathbb{F}_2} \mathrm{Sel}(T,\mathbb{1}_K)} \prod_{v \in \Sigma} (-1)^{h_v(\mathbb{1}_K, \gamma_v)},$$

and the right-hand side depends only on γ. In particular, by Theorem 9.4 we have

$$\lim_{X\to\infty} \frac{|\{\chi \in \mathcal{C}(K, X) : \chi|_\Gamma = \gamma \text{ and } \dim_{\mathbb{F}_2} \mathrm{Sel}(T, \chi) \text{ is even}\}|}{|\{\chi \in \mathcal{C}(K, X) : \chi|_\Gamma = \gamma\}|} = \frac{1}{2},$$

and if $p = 2$ then this is in fact an equality for all sufficiently large X rather than a limit. Averaging over all $\gamma \in \Gamma$ gives the result (note that the sets $\{\chi \in \mathcal{C}(K, X) : \chi|_\Gamma = \gamma\}$ all have the same size for sufficiently

large X as the restriction map $\chi \mapsto \chi|_\Gamma$ is a homomorphism and is surjective for X sufficiently large by Lemma 7.2). □

We now turn to the proof of Theorem 9.4.

Definition 9.6. Fix an $\mathbb{F}_p$-basis $\{\phi_1, \dots, \phi_r\}$ for $H^1(K(T)/K, \boldsymbol{\mu}_p)$. Further, define the homomorphism $f : \Gamma \to \boldsymbol{\mu}_p^r$ by setting

$$f(\gamma) = ([\phi_i, \gamma]_\Sigma)_{i=1}^r$$

where here we view the ϕ_i inside $\prod_{v\in\Sigma} H^1(K_v, \boldsymbol{\mu}_p)$ via the product of the natural restriction maps, and $[\cdot,\cdot]_\Sigma$ is the pairing of Notation 8.5 (we allow the case $r = 0$ in which case $\boldsymbol{\mu}_p^r$ is the trivial group).

Remark 9.7. Since we have taken Σ large enough that the map

$$H^1(K(T)/K, \boldsymbol{\mu}_p) \to \prod_{v\in\Sigma} H^1(K_v, \boldsymbol{\mu}_p)$$

is injective, it follows from the nondegeneracy of the pairing $[\cdot,\cdot]_\Sigma$ that f is surjective.

Definition 9.8. For each $n \geq 1$ and $\eta \in \boldsymbol{\mu}_p^r$, define

$$t_X(\eta) = \frac{|\{\chi \in \mathcal{C}(K, X) : f(\chi|_\Gamma) = \eta,\ w(\chi) = 1\}|}{|\{\chi \in \mathcal{C}(K, X) : f(\chi|_\Gamma) = \eta\}|}.$$

The following lemma reduces the problem of understanding $s_X(\gamma)$ as γ ranges over the elements of Γ, to understanding $t_X(\eta)$ as η ranges over the elements of $\boldsymbol{\mu}_p^r$.

Lemma 9.9 [Klagsbrun et al. 2013, Lemma 8.4]. *Let $\gamma \in \Gamma$. Then for X sufficiently large we have*

$$s_X(\gamma) = t_X(f(\gamma)).$$

Proof. Let P denote the set of primes of K not in Σ and which split completely in $K(T)/K$, and let $\gamma' \in \Gamma$ be such that $f(\gamma') = f(\gamma)$. Then $\gamma'\gamma^{-1}$ is in the kernel of f so by Lemma 8.6 there is $\chi_{\gamma,\gamma'} \in \mathcal{C}(K)$ with $\chi_{\gamma,\gamma'}|_\Gamma = \gamma'\gamma^{-1}$ and such that $\chi_{\gamma,\gamma'}$ is unramified outside $\Sigma \cup P$. Now for any $\chi \in \mathcal{C}(K)$ we have $w(\chi) = w(\chi\chi_{\gamma,\gamma'})$ since the sets of primes not in Σ where χ and $\chi_{\gamma,\gamma'}$ ramify differ only at primes $\mathfrak{p} \in P$, and at such primes we have

$$\epsilon(\mathrm{Frob}_\mathfrak{p}) = \epsilon(1) = (-1)^{\dim T} = 1$$

(where as usual $\mathrm{Frob}_\mathfrak{p}$ denotes the Frobenius element at $\mathfrak{p}$ in $K(T)/K$). Thus if X is sufficiently large that $\chi_{\gamma,\gamma'}$ is in $\mathcal{C}(K, X)$, multiplication by $\chi_{\gamma,\gamma'}$ gives a bijection between the set

$$\{\chi \in \mathcal{C}(K, X) : \chi|_\Gamma = \gamma,\ w(\chi) = 1\}$$

and the set

$$\{\chi \in \mathcal{C}(K, X) : \chi|_\Gamma = \gamma',\ w(\chi) = 1\},$$

as well as between the same two sets with the conditions on $w(\chi)$ removed.

Writing $\eta = f(\gamma)$, it follows that for X sufficiently large we have

$$t_X(\eta) = \frac{\sum_{\gamma' \in f^{-1}(\{\eta\})} |\{\chi \in \mathcal{C}(K,X) : \chi|_\Gamma = \gamma',\ w(\chi) = 1\}|}{\sum_{\gamma' \in f^{-1}(\{\eta\})} |\{\chi \in \mathcal{C}(K,X) : \chi|_\Gamma = \gamma'\}|}$$
$$= \frac{|f^{-1}(\{\eta\})| \cdot |\{\chi \in \mathcal{C}(K,X) : \chi|_\Gamma = \gamma,\ w(\chi) = 1\}|}{|f^{-1}(\{\eta\})| \cdot |\{\chi \in \mathcal{C}(K,X) : \chi|_\Gamma = \gamma\}|} = s_X(\gamma)$$

as desired. □

We now study the quantities $t_X(\eta)$ as η ranges over $\boldsymbol{\mu}_p^r$, splitting into cases according to $p = 2$ or $p > 2$.

9A. ***The case where $p = 2$ and ϵ fails to be a homomorphism.*** Suppose now that $p = 2$ and ϵ fails to be a homomorphism.

Definition 9.10. Define the map $\theta : \mathcal{C}(K) \to \boldsymbol{\mu}_2^r \times \{\pm 1\}$ by setting

$$\theta(\chi) := (f(\chi|_\Gamma), w(\chi)).$$

The following observation will be crucial to our method. We remark that it fails for $p > 2$.

Lemma 9.11. *The map θ is a homomorphism.*

Proof. Since both the restriction map $\mathcal{C}(K) \to \Gamma$ and the map $f : \Gamma \to \boldsymbol{\mu}_2^r$ are homomorphisms, it suffices to show that $w : \mathcal{C}(K) \to \{\pm 1\}$ is a homomorphism.

For each $v \notin \Sigma$, define a map $w_v : \mathcal{C}(K_v) \to \{\pm 1\}$ by

$$w_v(\chi) = \begin{cases} (-1)^{\dim_{\mathbb{F}_2} T^{G_{K_v}}} & \chi \text{ ramified,} \\ 1 & \text{else.} \end{cases}$$

Since w is the product of the w_v over $v \notin \Sigma$, it suffices to show that each w_v is a homomorphism. To see this, note that as $v \notin \Sigma$, K_v has odd residue characteristic. In particular, the product of any two ramified characters of K_v is unramified, and the product of a ramified character with an unramified character is again ramified. □

Remark 9.12. For $X > 0$ write θ_X for the restriction of θ to $\mathcal{C}(K,X)$. Then for each $\eta \in \boldsymbol{\mu}_2^r$ we have

$$t_X(\eta) = \frac{|\theta_X^{-1}((\eta,1))|}{|\theta_X^{-1}((\eta,1))| + |\theta_X^{-1}((\eta,-1))|}.$$

Now (for $X > 1$), $\mathcal{C}(K,X)$ is a group and θ_X is a homomorphism. Thus the fibers over points in the image of θ_X have the same size, being cosets of the kernel. In light of Lemma 9.9, Theorem 9.4(i) is equivalent to the statement that, if ϵ fails to be a homomorphism, then θ_X is surjective for sufficiently large $X > 0$. Since $\boldsymbol{\mu}_2^r \times \{\pm 1\}$ is a finite group this is, in turn, equivalent to the statement that if ϵ fails to be a homomorphism then θ is surjective. This is the statement we now study, and prove in Proposition 9.14.

We now fix a collection of global characters $\{\varphi(\mathfrak{p})\}_{\mathfrak{p} \notin \Sigma}$ satisfying (8.11). Each $\varphi(\mathfrak{p})$ is ramified at $\mathfrak{p}$, yet unramified outside $\Sigma \cup \{\mathfrak{p}\}$. Lemma 8.12 allows us to evaluate the map θ on the $\varphi(\mathfrak{p})$.

Lemma 9.13. *For each* $\mathfrak{p} \notin \Sigma$ *we have*

$$\theta(\varphi(\mathfrak{p})) = ((\phi_i(\mathrm{Frob}_\mathfrak{p}))_{i=1}^r, \epsilon(\mathrm{Frob}_\mathfrak{p}))$$

where here $\mathrm{Frob}_\mathfrak{p} \in G$ *denotes the Frobenius element at* $\mathfrak{p}$ *in* $K(T)/K$.

Proof. Since amongst the primes not in Σ the character $\varphi(\mathfrak{p})$ only ramifies at $\mathfrak{p}$, we have $w(\varphi(\mathfrak{p})) = \epsilon(\mathrm{Frob}_\mathfrak{p})$ by definition. We have $f(\varphi(\mathfrak{p})|_\Gamma) = (\phi_i(\mathrm{Frob}_\mathfrak{p}))_{i=1}^r$ by Lemma 8.12. □

Proposition 9.14. *The map* $\theta : \mathcal{C}(K) \to \boldsymbol{\mu}_2^r \times \{\pm 1\}$ *is surjective if and only if* ϵ *fails to be a homomorphism.*

Proof. Note that the subgroup $\mathcal{U}$ of $\mathcal{C}(K)$ consisting of characters unramified outside Σ is in the kernel of θ, and the quotient $\mathcal{C}(K)/\mathcal{U}$ is generated by the $\varphi(\mathfrak{p})$ as $\mathfrak{p}$ ranges over primes not in Σ.

By the Chebotarev density theorem, each conjugacy class in $G = \mathrm{Gal}(K(T)/K)$ arises as $\mathrm{Frob}_\mathfrak{p}$ for some $\mathfrak{p} \notin \Sigma$ and so by Lemma 9.13 it follows that the image of θ is the subgroup of $\boldsymbol{\mu}_2^r \times \{\pm 1\}$ generated by the set

$$\{((\phi_i(\sigma))_{i=1}^r, \epsilon(\sigma)) : \sigma \in G\}$$

(note that for $\sigma \in G$, both $\epsilon(\sigma)$ and the $\phi_i(\sigma)$ depend only on the conjugacy class of σ in G).

Recall that the set $\{\phi_i : 1 \le i \le r\}$ is a basis for $H^1(K(T)/K, \boldsymbol{\mu}_2) = \mathrm{Hom}(G, \boldsymbol{\mu}_2)$. To make this more explicit denote by G^2 the subgroup of G generated by the squares of all the elements of G. It's a normal subgroup and the quotient G/G^2 is an abelian group of exponent 2. That is, G/G^2 is a finite dimensional $\mathbb{F}_2$-vector space. Since every homomorphism from G to $\boldsymbol{\mu}_2$ factors through G/G^2 we have

$$\mathrm{Hom}(G, \boldsymbol{\mu}_2) = \mathrm{Hom}(G/G^2, \boldsymbol{\mu}_2)$$

and the right-hand group is just the dual of G/G^2 as an $\mathbb{F}_2$-vector space. In particular, the map $G/G^2 \to \boldsymbol{\mu}_2^r$ sending σ to $(\phi_i(\sigma))_{i=1}^r$ is an isomorphism.

Combining the above we arrive at a purely group theoretic criterion: θ is surjective if and only if the set

$$S := \{(\bar{\sigma}, \epsilon(\sigma)) : \sigma \in G\}$$

generates $G/G^2 \times \{\pm 1\}$, where here for $\sigma \in G$ we write $\bar{\sigma}$ for the image of σ in G/G^2.

Suppose now that ϵ is a homomorphism. Then ϵ necessarily factors through G/G^2 and we see that S generates an index 2 subgroup of $G/G^2 \times \{\pm 1\}$, so that θ is not surjective in this case.

Conversely, suppose that ϵ fails to be a homomorphism and write H for the subgroup of $G/G^2 \times \{\pm 1\}$ generated by S. By assumption, we may find σ and τ in G with $\epsilon(\sigma\tau) = -\epsilon(\sigma)\epsilon(\tau)$. Then

$$(\bar{\sigma}, \epsilon(\sigma)) \cdot (\bar{\tau}, \epsilon(\tau)) \cdot (\overline{\sigma\tau}, \epsilon(\sigma\tau)) = (\overline{(\sigma\tau)^2}, -1) = (1, -1)$$

is in H (here the first 1 denotes the identity in G/G^2). Then for any $\sigma \in G$, both $(\bar{\sigma}, \epsilon(\sigma))$ and

$$(\bar{\sigma}, -\epsilon(\sigma)) = (1, -1) \cdot (\bar{\sigma}, \epsilon(\sigma))$$

are in H. Thus $H = G/G^2 \times \{\pm 1\}$ and θ is surjective. □

Proof of Theorem 9.4(i). By Remark 9.12 we see that Theorem 9.4(i) holds if and only if θ is surjective whenever ϵ fails to be a homomorphism. The result now follows from Proposition 9.14. □

9B. *The case where $p > 2$ and ϵ is nontrivial when restricted to* $\mathbf{Gal}(K(T)/K(\boldsymbol{\mu}_p))$. Suppose now that $p > 2$ and that the restriction of ϵ to $\mathrm{Gal}(K(T)/K(\boldsymbol{\mu}_p))$ is nontrivial.

We begin by defining a slight refinement of the quantity $t_X(\eta)$.

Definition 9.15. Fix an enumeration of the primes $\mathfrak{p} \notin \Sigma$ such that if $i \leq j$ then $N(\mathfrak{p}_i) \leq N(\mathfrak{p}_j)$, and for each $n \geq 1$ define the subgroup $\mathcal{C}_n(K)$ of $\mathcal{C}(K)$ by

$$\mathcal{C}_n(K) := \{\chi \in \mathcal{C}(K) : \chi \text{ is unramified outside } \Sigma \cup \{\mathfrak{p}_1, \dots, \mathfrak{p}_n\}\}.$$

Further, for each $n \geq 1$ and $\eta \in \boldsymbol{\mu}_p^r$, define

$$\hat{t}_n(\eta) := \frac{|\{\chi \in \mathcal{C}_n(K) : f(\chi|_\Gamma) = \eta, w(\chi) = 1\}|}{|\{\chi \in \mathcal{C}_n(K) : f(\chi|_\Gamma) = \eta\}|} - \frac{1}{2}.$$

Remark 9.16. Note that we subtract $\frac{1}{2}$ in the definition of $\hat{t}_n(\eta)$ whilst we did not in the definition of $t_X(\eta)$. This will neaten the statement of some results in the rest of the section. Clearly for any $\eta \in \boldsymbol{\mu}_p^r$, to show that $\lim_{X\to\infty} t_X(\eta) = \frac{1}{2}$ it suffices to show that $\lim_{n\to\infty} \hat{t}_n(\eta) = 0$.

As in the case $p = 2$ we now fix a collection of global characters $\{\varphi(\mathfrak{p})\}_{\mathfrak{p}\notin\Sigma, \boldsymbol{\mu}_p \subseteq K_\mathfrak{p}}$ satisfying (8.11).

Lemma 9.17. *Fix $n \geq 1$. Then if $\boldsymbol{\mu}_p \subsetneq K_{\mathfrak{p}_{n+1}}$ we have $\mathcal{C}_{n+1}(K) = \mathcal{C}_n(K)$. On the other hand, if $\boldsymbol{\mu}_p \subseteq K_{\mathfrak{p}_{n+1}}$ then we have*

$$\mathcal{C}_{n+1}(K) = \bigsqcup_{i=0}^{p-1} \varphi(\mathfrak{p}_{n+1})^i \cdot \mathcal{C}_n(K).$$

Proof. In each case this follows from the structure of $\mathcal{C}(K_{\mathfrak{p}_{n+1}})$; see [Klagsbrun et al. 2013, Lemma 8.3]. □

Definition 9.18. Let V be the regular representation of $\boldsymbol{\mu}_p^r$ over $\mathbb{C}$, so that V has basis $\{e_\eta : \eta \in \boldsymbol{\mu}_p^r\}$ on which $\boldsymbol{\mu}_p^r$ acts via $\eta' \cdot e_\eta = e_{\eta'\eta}$. For each $n \geq 1$ define

$$\hat{\boldsymbol{t}}_n := \sum_{\eta\in\boldsymbol{\mu}_p^r} \hat{t}_n(\eta) e_\eta \in V.$$

Further, for $\sigma \in \mathrm{Gal}(K(T)/K(\boldsymbol{\mu}_p))$, define $\rho(\sigma) := (\phi_i(\sigma))_{i=1}^r \in \boldsymbol{\mu}_p^r$ and

$$M(\sigma) := \tfrac{1}{p}\left(1 + \epsilon(\sigma) \sum_{i=1}^{p-1} \rho(\sigma)^i\right) \in \mathrm{End}(V).$$

Remark 9.19. For $\sigma \in \mathrm{Gal}(K(T)/K(\boldsymbol{\mu}_p))$ the element $M(\sigma)$ depends only on the conjugacy class of σ in G. Indeed, for each $1 \leq i \leq r$ and $g \in G$, the cocycle relation for ϕ_i gives $\phi_i(g\sigma g^{-1}) = g\phi_i(\sigma)$. It now follows that for each i, $\sum_{j=1}^{p-1} \phi_i(\sigma)^j$ depends only on the conjugacy class of σ in G. Since the same is true for $\epsilon(\sigma)$ we are done.

Lemma 9.20. *Fix $n \geq 1$. If $\boldsymbol{\mu}_p \subsetneq K_{\mathfrak{p}_{n+1}}$ then we have $\hat{\boldsymbol{t}}_{n+1} = \hat{\boldsymbol{t}}_n$. On the other hand, if $\boldsymbol{\mu}_p \subseteq K_{\mathfrak{p}_{n+1}}$ then we have the following recurrence relation for $\hat{\boldsymbol{t}}_n$:*

$$\hat{\boldsymbol{t}}_{n+1} = M(\mathrm{Frob}_{\mathfrak{p}_{n+1}})\hat{\boldsymbol{t}}_n,$$

where here $\mathrm{Frob}_{\mathfrak{p}_{n+1}} \in G$ denotes the Frobenius element at $\mathfrak{p}_{n+1}$ in $K(T)/K$.

Proof. If $\boldsymbol{\mu}_p \subsetneq K_{\mathfrak{p}_{n+1}}$ then $\mathcal{C}_{n+1}(K) = \mathcal{C}_n(K)$ and the result is clear.

Suppose now that $\boldsymbol{\mu}_p \subseteq K_{\mathfrak{p}_{n+1}}$ and define the map $\theta : \mathcal{C}(K) \to \boldsymbol{\mu}_p^r \times \{\pm 1\}$ by

$$\theta(\chi) := (f(\chi|_\Gamma), w(\chi))$$

(note that, unlike the case $p = 2$ this is not a homomorphism). Then Lemma 8.12 gives

$$\theta(\varphi(\mathfrak{p}_n)) = (\rho(\mathrm{Frob}_{\mathfrak{p}_n}), \epsilon(\mathrm{Frob}_{\mathfrak{p}_n})).$$

Moreover, if $\chi_0 \in \mathcal{C}_n(K)$ then we have

$$\theta(\chi_0 \cdot \varphi(\mathfrak{p}_{n+1})^i) = \theta(\chi_0) \cdot \theta(\varphi(\mathfrak{p}_{n+1})^i)$$

since the sets of primes not in Σ at which χ_0 and $\varphi(\mathfrak{p}_{n+1})^i$ ramify are disjoint. Writing σ for $\mathrm{Frob}_{\mathfrak{p}_{n+1}}$, this gives

$$\theta(\chi_0 \cdot \varphi(\mathfrak{p}_{n+1})^i) = \begin{cases} \theta(\chi_0) & i = 0, \\ \theta(\chi_0) \cdot (\rho(\sigma)^i, \epsilon(\sigma)) & 1 \leq i \leq p-1. \end{cases}$$

It now follows from Lemma 9.17 that for each $\eta \in \boldsymbol{\mu}_p^r$ we have

$$\begin{aligned} &|\{\chi \in \mathcal{C}_{n+1}(K) : \theta(\chi) = (\eta, 1)\}| \\ &\qquad = \sum_{i=0}^{p-1} |\{\chi \in \varphi(\mathfrak{p}_{n+1})^i \cdot \mathcal{C}_n(K) : \theta(\chi) = (\eta, 1)\}| \\ &\qquad = |\{\chi_0 \in \mathcal{C}_n(K) : \theta(\chi_0) = (\eta, 1)\}| + \sum_{i=1}^{p-1} |\{\chi_0 \in \mathcal{C}_n(K) : \theta(\chi_0) = (\eta \cdot \rho(\sigma)^{-i}, \epsilon(\sigma))\}|. \end{aligned}$$

Dividing through by $|\mathcal{C}_{n+1}(K)| = p|\mathcal{C}_n(K)|$ gives

$$\hat{t}_{n+1}(\eta) = \tfrac{1}{p}\left(\hat{t}_n(\eta) + \epsilon(\sigma) \sum_{i=1}^{p-1} \hat{t}_n(\rho(\sigma)^{-i} \cdot \eta)\right)$$

and the result now follows from the definition of $M(\sigma)$. □

Lemma 9.21. *For any $m \geq 1$ and $\sigma \in \mathrm{Gal}(K(T)/K(\boldsymbol{\mu}_p))$ we have*

$$M(\sigma)^m = \begin{cases} M(\sigma) & \epsilon(\sigma) = 1, \\ \left(\frac{2}{p}\left(\frac{2-p}{p}\right)^m - \frac{2-p}{p}\left(\frac{2}{p}\right)^m\right) \mathrm{id}_V + \left(\left(\frac{2}{p}\right)^m - \left(\frac{2-p}{p}\right)^m\right) M(\sigma) & \epsilon(\sigma) = -1. \end{cases}$$

In particular, for $p > 2$ and writing $\|\cdot\|$ for the operator norm on $\mathrm{End}(V)$*, we have*

$$\lim_{m\to\infty} \|M(\sigma)^m\| = \begin{cases} \|M(\sigma)\| & \epsilon(\sigma) = 1, \\ 0 & \epsilon(\sigma) = -1. \end{cases}$$

Proof. Fix $\sigma \in \mathrm{Gal}(K(T)/K(\boldsymbol{\mu}_p))$ and define

$$T(\sigma) := \tfrac{1}{p} \sum_{i=0}^{p-1} \rho(\sigma)^i.$$

Then $T(\sigma)$ is an idempotent in $\mathrm{End}(V)$ (e.g., by orthogonality of characters of $\boldsymbol{\mu}_p$ or by explicit computation) so that $T(\sigma)^m = T(\sigma)$ for each $m \geq 1$. Note that we have

$$M(\sigma) = \begin{cases} T(\sigma) & \epsilon(\sigma) = 1, \\ \frac{2}{p} - T(\sigma) & \epsilon(\sigma) = -1. \end{cases}$$

If $\epsilon(\sigma) = 1$ this immediately gives $M(\sigma)^m = M(\sigma)$, whilst if $\epsilon(\sigma) = -1$ the result now follows easily either by induction on m or by expanding $\left(\frac{2}{p} - T(\sigma)\right)^m$ with the binomial theorem.

Since $p > 2$ we have both

$$\lim_{m\to\infty} \left(\frac{2}{p}\right)^m = 0 \quad \text{and} \quad \lim_{m\to\infty} \left(\frac{2-p}{p}\right)^m = 0,$$

from which the statement about $\lim_{m\to\infty} \|M(\sigma)^m\|$ follows immediately. □

Proposition 9.22. *Suppose that $p > 2$ and ϵ is nontrivial when restricted to* $\mathrm{Gal}(K(T)/K(\boldsymbol{\mu}_p))$*. Then for each $\eta \in \boldsymbol{\mu}_p^r$ we have*

$$\lim_{n\to\infty} \hat{t}_n(\eta) = 0.$$

Proof. Write $H := \mathrm{Gal}(K(T)/K(\boldsymbol{\mu}_p))$, and note that this is a normal subgroup of G. For each $n \geq 1$ we have $\boldsymbol{\mu}_p \subseteq K_{\mathfrak{p}_n}$ if and only if $\mathrm{Frob}_{\mathfrak{p}_n} \in H$. By Lemma 9.20, for each $n \geq 1$ we have

$$\hat{\boldsymbol{t}}_n = \left(\prod_{\substack{i=2 \\ \mathrm{Frob}_{\mathfrak{p}_i} \in H}}^{n} M(\mathrm{Frob}_{\mathfrak{p}_i}) \right) \hat{\boldsymbol{t}}_1. \tag{9.23}$$

Write $C_1, \ldots, C_l$ for the conjugacy classes in G that are contained in H and, for each i, fix a representative σ_i for C_i. Further, for each $1 \leq i \leq l$, define

$$m_i(n) := |\{2 \leq j \leq n : \mathrm{Frob}_{\mathfrak{p}_j} \in C_i\}|.$$

Since the group ring $\mathbb{C}[\boldsymbol{\mu}_p^r]$ is commutative, the matrices $M(\sigma_i)$ all mutually commute and we may group like terms in (9.23) to obtain (cf. Remark 9.19)

$$\hat{\boldsymbol{t}}_n = \left(\prod_{i=1}^{l} M(\sigma_i)^{m_i(n)} \right) \hat{\boldsymbol{t}}_1.$$

Writing $\|\cdot\|$ for the usual Euclidean norm on V (with respect to the basis $\{e_\eta : \eta \in \boldsymbol{\mu}_p^r\}$), we have

$$\|\hat{\boldsymbol{t}}_n\| = \left\| \left(\prod_{i=1}^{l} M(\sigma_i)^{m_i(n)} \right) \hat{\boldsymbol{t}}_1 \right\| \leq \left(\prod_{i=1}^{l} \|M(\sigma_i)\|^{m_i(n)} \right) \|\hat{\boldsymbol{t}}_1\|.$$

By the Chebotarev density theorem each of the $m_i(n)$ tend to infinity with n and, since we have assumed there is at least one i with $\epsilon(\sigma_i) = -1$, it follows from Lemma 9.21 that

$$\lim_{n\to\infty} \|\hat{\boldsymbol{t}}_n\| = 0.$$

That is, $\lim_{n\to\infty} \hat{t}_n(\eta) = 0$ for each $\eta \in \boldsymbol{\mu}_p^r$. □

Proof of Theorem 9.4(ii). Fix $\gamma \in \Gamma$ and write $\eta = f(\gamma)$. Then by Lemma 9.9, for all X sufficiently large we have $s_X(\gamma) = t_X(\eta)$. It follows from Proposition 9.22 that $\lim_{X\to\infty} t_X(\eta) = \frac{1}{2}$, from which the result follows. □

10. Twisting data for abelian varieties ($p = 2$)

In this section let K be a number field and $(A/K, \lambda)$ a principally polarized abelian variety. In the notation of Sections 6–9 we take $p = 2$ and $T = A[2]$ endowed with the Weil pairing $(\cdot, \cdot)_\lambda$. Let Σ be a finite set of places of K containing all archimedean places, all places dividing 2, and all places at which A has bad reduction. Then T is unramified outside Σ.

We now endow T with a global metabolic structure and twisting data in such a way that for $\chi \in \mathcal{C}(K)$ the associated Selmer group $\mathrm{Sel}(A[2], \chi)$ agrees with the 2-Selmer group $\mathrm{Sel}_2(A^\chi/K)$ of the quadratic twist of A by χ. For elliptic curves this is done in [Klagsbrun et al. 2013, §5]. Our definition of the global metabolic structure and twisting data will be a direct generalization of theirs. The main difficulty is establishing Lemma 10.6 which for elliptic curves is [Klagsbrun et al. 2013, Lemma 5.2(ii)] and for Jacobians of odd degree hyperelliptic curves is [Yu 2016, Theorem 5.10]. We will deduce the general case from the results of Section 4E concerning the behavior of certain Theta groups under quadratic twist.

10A. ***A global metabolic structure on A[2].*** For a place v of K write

$$\delta_v : A(K_v)/2A(K_v) \hookrightarrow H^1(K_v, A[2])$$

for the connecting homomorphism in the multiplication-by-2 Kummer sequence.

Definition 10.1. Let $\mathcal{P}$ denote the Poincaré line bundle on $A \times A^\vee$. For each place v of K write $\mathcal{L}_v$ for the pull back of $\mathcal{L} = (1, \lambda)^*\mathcal{P}$ to a line bundle on A/K_v and let $\mathcal{G}(\mathcal{L}_v)$ denote the associated Theta group. Then we define $q_{A,\lambda,v}$ to be the map

$$q_{A,\lambda,v} : H^1(K_v, A[2]) \to H^2(K_v, \overline{K_v}^\times) = \mathrm{Br}(K_v) \xrightarrow{\mathrm{inv}_v} \mathbb{Q}/\mathbb{Z}$$

where the first map is the connecting map associated to the short exact sequence of G_{K_v}-modules

$$0 \to \overline{K_v}^\times \to \mathcal{G}(\mathcal{L}_v) \to A[2] \to 0 \tag{10.2}$$

of Remark 4.14.

Lemma 10.3. *Let v be a place of K. Then*:

(i) *$q_{A,\lambda,v}$ is a quadratic form on $H^1(K_v, A[2])$ whose associated bilinear pairing is the local Tate pairing corresponding to $(\cdot,\cdot)_\lambda$.*

(ii) *The image of $A(K_v)/2A(K_v)$ under δ_v is a Lagrangian subspace of $H^1(K_v, A[2])$ with respect to $q_{A,\lambda,v}$.*

In particular, $q_{A,\lambda,v}$ is a Tate quadratic form on $H^1(K_v, A[2])$ in the sense of Definition 6.1.

Proof. Part (i) is [Poonen and Rains 2012, Corollary 4.7] whilst Proposition 4.9 of op. cit. gives (ii). □

Remark 10.4. In contrast to the case of elliptic curves the quadratic form $q_{A,\lambda,v}$ in general takes values in $\frac{1}{4}\mathbb{Z}/\mathbb{Z}$ rather than just $\frac{1}{2}\mathbb{Z}/\mathbb{Z}$, which is the reason for allowing $\mathbb{Q}/\mathbb{Z}$-valued quadratic forms in Definition 6.1 rather than just those valued in $\mathbb{F}_2$. See [Poonen and Rains 2012, Remark 4.16] for an example of this phenomenon.

Corollary 10.5. *The collection $\boldsymbol{q} = (q_{A,\lambda,v})_v$ defines a global metabolic structure on $A[2]$.*

Proof. By Lemma 10.3(ii) $q_{A,\lambda,v}$ admits a Lagrangian subspace making $(H^1(K_v, T), q_{A,\lambda,v})$ into a metabolic space for each place v of K. Moreover, if $v \notin \Sigma$ then $\mathrm{im}(\delta_v) = H^1_{\mathrm{ur}}(K_v, A[2])$ (see e.g., [Poonen and Rains 2012, Proposition 4.12] and the preceding remark). In particular, by Lemma 10.3(ii), $q_{A,\lambda,v}$ is unramified at each such place.

Finally, let $\mathfrak{a} \in H^1(K, A[2])$. Write

$$q : H^1(K, A[2]) \to H^2(K, \bar{K}^\times) = \mathrm{Br}(K)$$

for the connecting homomorphism associated to the sequence (10.2) viewed over K instead of K_v (with $\mathcal{L}_v$ replaced by $\mathcal{L} := (1,\lambda)^*\mathcal{P}$). Then $q(\mathfrak{a}) \in \mathrm{Br}(K)$ and we have

$$\sum_{v \in M_K} q_v(\mathfrak{a}_v) = \sum_{v \in M_K} \mathrm{inv}_v\, q(\mathfrak{a}) = 0,$$

the last equality following from reciprocity for the Brauer group of K. □

10B. ***Twisting data associated to A/K.*** We now define the twisting data $\boldsymbol{\alpha}$.

Fix a place v of K and $\chi \in \mathcal{C}(K_v)$, and let (A^χ, ψ) denote the quadratic twist of A by χ. By Lemma 4.16 $\lambda_\chi := (\psi^\vee)^{-1}\lambda\psi^{-1}$ is a principal polarization on A^χ, defined over K_v. In particular, associated to the pair (A^χ, λ_χ) we have a quadratic form $q_{A^\chi,\lambda_\chi,v}$ on $H^1(K_v, A^\chi[2])$.

Lemma 10.6. *The isomorphism $H^1(K_v, A^\chi[2]) \cong H^1(K_v, A[2])$ induced by ψ identifies the quadratic forms $q_{A,\lambda,v}$ and $q_{A^\chi,\lambda_\chi,v}$.*

Proof. Take the long exact sequences for Galois cohomology associated to the commutative diagram of Lemma 4.20. □

Definition 10.7. For $\chi \in \mathcal{C}(K_v)$ define $\alpha_v(\chi) \subseteq H^1(K_v, A[2])$ to be the image of the map

$$A^\chi(K_v)/2A^\chi(K_v) \hookrightarrow H^1(K_v, A^\chi[2]) \xrightarrow{\sim} H^1(K_v, A[2])$$

the first map arising from the multiplication-by-2 Kummer sequence for A^χ and the latter being induced by ψ^{-1}. Note that by combining Lemma 10.6 with Lemma 10.3(ii) applied to A^χ/K_v we see that $\alpha_v(\chi)$ is a Lagrangian subspace of $H^1(K_v, A[2])$.

As in Definition 6.11, for χ_1 and χ_2 elements of $\mathcal{C}(K_v)$ we set

$$h_v(\chi_1, \chi_2) = \dim_{\mathbb{F}_p}(\alpha_v(\chi_1)/(\alpha_v(\chi_1) \cap \alpha_v(\chi_2))).$$

Lemma 10.8. *For each quadratic character $\chi \in \mathcal{C}(K_v)$, let L_χ denote the extension of K_v cut out by χ. Then*

$$h_v(\mathbb{1}, \chi_v) = \dim_{\mathbb{F}_2} A(K_v)/N_{L_\chi/K_v} A(L_\chi)$$

where here $N_{L_\chi/K_v} : A(L_\chi) \to A(K_v)$ is the "local norm map" sending $P \in A(L_\chi)$ to

$$N_{L_\chi/K_v}(P) := \sum_{\sigma \in \mathrm{Gal}(L_\chi/K_v)} \sigma(P).$$

Proof. This is shown in [Mazur and Rubin 2007, Proposition 5.2]. Whilst that statement is for the case of elliptic curves and for twists by characters of order $p > 2$, the proof carries over unchanged to our case. See also [Kramer 1981, Proposition 7]. □

The following lemma evaluates the cokernel of the local norm map in certain cases.

Lemma 10.9. *Let v be a place of K and $\chi \in \mathcal{C}(K_v)$. As above, let L_χ denote the extension of K_v cut out by χ.*

(i) *Suppose $v \nmid 2$ is nonarchimedean and that A has good reduction at v. If χ is unramified then*

$$\dim_{\mathbb{F}_2} A(K_v)/N_{L_\chi/K_v} A(L_\chi) = 0.$$

On the other hand, if χ is ramified then $N_{L_\chi/K_v} A(L_\chi) = 2A(K_v)$ and, in particular, we have

$$\dim_{\mathbb{F}_2} A(K_v)/N_{L_\chi/K_v} A(L_\chi) = \dim_{\mathbb{F}_2} A(K_v)[2].$$

(ii) *Suppose v is archimedean and χ nontrivial. Then*

$$\dim_{\mathbb{F}_2} A(K_v)/N_{L_\chi/K_v} A(L_\chi) = \dim_{\mathbb{F}_2} A(K_v)[2] - g$$

where $g = \dim A$ is the dimension of A.

Proof.

(i) The case where χ is unramified is a result of Mazur [Mazur 1972, Corollary 4.4]. For χ ramified the case where A is an elliptic curve is [Mazur and Rubin 2007, Lemma 5.5(ii)] and the argument for general abelian varieties is identical.

(ii) By assumption L_χ/K_v is the extension $\mathbb{C}/\mathbb{R}$. Since A/K_v is an abelian variety of dimension g over the reals we have an isomorphism of real Lie groups

$$A(K_v) \cong (\mathbb{R}/\mathbb{Z})^g \times (\mathbb{Z}/2\mathbb{Z})^m \tag{10.10}$$

for some $0 \leq m \leq g$ (see, for example, [Silhol 1989, Proposition 1.9 and Remark.12]). Now N_{L_χ/K_v} is a continuous map from the connected group $A(L_\chi)$ to $A(K_v)$ (for the complex and real topologies respectively) and it follows that the image of N_{L_χ/K_v} is contained in the connected component of the identity in $A(K_v)$, which we denote $A^0(K_v)$. Under the isomorphism (10.10), $A^0(K_v)$ is the factor corresponding to $(\mathbb{R}/\mathbb{Z})^g$. On the other hand, we have $2A(K_v) \subseteq N_{L_\chi/K_v}A(L_\chi)$ and we see again from (10.10) that multiplication by 2 is surjective on $A^0(K_v)$. Thus $N_{L_\chi/K_v}A(L_\chi) = A^0(K_v)$. Appealing to (10.10) one last time we obtain $|A(K_v)/N_{L_\chi/K_v}A(L_\chi)| = 2^{-g}|A(K_v)[2]|$. □

Proposition 10.11. *The collection of maps $\boldsymbol{\alpha} = (\alpha_v)_v$ defines twisting data with respect to $(A[2], \boldsymbol{q}, \Sigma)$. Moreover, we have*

$$\mathrm{Sel}(A[2], \chi) \cong \mathrm{Sel}_2(A^\chi/K)$$

where $\mathrm{Sel}(A[2], \chi)$ is defined with respect to $(A[2], \boldsymbol{q}, \Sigma, \boldsymbol{\alpha})$ as in Definition 6.10.

Proof. Note that since $p = 2$ the group $\mathcal{F}(K_v)$ appearing in the definition of twisting data (Definition 6.8) is equal to $\mathcal{C}(K_v)$. For each place v of K and $\chi_v \in \mathcal{C}(K_v)$, the subspace $\alpha_v(\chi_v)$ of $H^1(K_v, A[2])$ is Lagrangian by Lemmas 10.6 and 10.3(ii) applied to A^χ/K_v. Moreover, if $v \notin \Sigma$ and χ_v is ramified then $\alpha_v(\chi_v)$ is an element of $\mathcal{H}_{\mathrm{ram}}(q_v)$. Indeed, by definition we need to show that $\alpha_v(\chi_v) \cap H^1_{\mathrm{ur}}(K_v, A[2]) = 0$. As before, as $v \notin \Sigma$ we have

$$H^1_{\mathrm{ur}}(K_v, A[2]) = \delta_v(A(K_v)/2A(K_v)) = \alpha_v(\mathbb{1}_v).$$

Combining Lemma 10.8 with Lemma 10.9 gives

$$\dim_{\mathbb{F}_2}(\alpha(\mathbb{1}_v)/\alpha_v(\chi_v) \cap \alpha_v(\mathbb{1}_v)) = \dim_{\mathbb{F}_2} A(K_v)/2A(K_v) = \dim_{\mathbb{F}_2} \alpha(\mathbb{1}_v)$$

whence $\alpha_v(\chi_v) \cap \alpha_v(\mathbb{1}_v) = 0$ as desired. Thus $\boldsymbol{\alpha}$ defines twisting data.

Finally, we will show that for $\chi \in \mathcal{C}(K)$ the associated Selmer group $\mathrm{Sel}(A[2], \chi)$ agrees with the classical Selmer group $\mathrm{Sel}_2(A^\chi/K)$. By the definition of $\mathrm{Sel}_2(A^\chi/K)$ and the maps α_v we have

$$\mathrm{Sel}_2(A^\chi/K) = \{\mathfrak{a} \in H^1(K, A[2]) : \mathfrak{a}_v \in \alpha_v(\chi_v) \text{ for all } v \in M_K\}.$$

On the other hand, we have

$$\mathrm{Sel}(A[2], \chi) = \{\mathfrak{a} \in H^1(K, A[2]) : \mathfrak{a}_v \in H^1_{\mathcal{S}(\chi)}(K_v, A[2]) \text{ for all } v \in M_K\}$$

where, as in Definition 6.10, $H^1_{\mathcal{S}(\chi)}(K_v, A[2]) = \alpha(\chi_v)$ if $v \in \Sigma$ or χ_v is ramified at v, and is equal to $H^1_{\mathrm{ur}}(K_v, A[2])$ otherwise.

In particular, to show that $\mathrm{Sel}(A[2], \chi) = \mathrm{Sel}_2(A^\chi/K)$ it suffices to show that $\alpha(\chi_v) = H^1_{\mathrm{ur}}(K_v, A[2])$ whenever $v \notin \Sigma$ and χ_v is unramified. But for such places we have $\alpha(\mathbb{1}_v) = H^1_{\mathrm{ur}}(K_v, A[2])$ and since χ_v is

unramified Lemma 10.9(i) gives $h(\mathbb{1}_v, \chi_v) = 0$. It now follows immediately that $\alpha(\chi_v) = H^1_{\mathrm{ur}}(K_v, A[2])$ as desired. □

10C. ***Main theorems for 2-Selmer ranks.*** Having interpreted the groups $\mathrm{Sel}_2(A^\chi/K)$ as those arising from twisting data we apply the results of the previous sections to deduce results about abelian varieties.

The following generalizes a theorem of Kramer [1981, Theorem 1] for elliptic curves and Yu [2016, Theorem 5.11] for odd degree hyperelliptic curves.

Theorem 10.12. *Let K be a number field, χ a quadratic character of K corresponding to the extension L/K, and A/K a principally polarized abelian variety. Then*

$$\dim_{\mathbb{F}_2} \mathrm{Sel}_2(A^\chi/K) \equiv \dim_{\mathbb{F}_2} \mathrm{Sel}_2(A/K) + \sum_{v \in M_K} \dim_{\mathbb{F}_2} A(K_v)/N_{L_w/K_v} A(L_w) \pmod 2$$

(*here w denotes any place of L extending v*).

Proof. Combine Theorem 6.12, Proposition 10.11 and Lemma 10.8. □

Theorem 10.13 (Theorem 1.1). *Let K be a number field, A/K a principally polarized abelian variety, and Σ the set consisting of all archimedean places of K, all places of bad reduction for A, and all places dividing 2. Define $\epsilon : \mathrm{Gal}(K(A[2])/K) \to \{\pm 1\}$ by $\sigma \mapsto (-1)^{\dim_{\mathbb{F}_2} A[2]^\sigma}$.*

(i) *If ϵ fails to be a homomorphism then for all sufficiently large X*

$$\frac{|\{\chi \in \mathcal{C}(K, X) : \dim_{\mathbb{F}_2} \mathrm{Sel}_2(A^\chi/K) \text{ is even}\}|}{|\mathcal{C}(K, X)|} = \frac{1}{2}.$$

(ii) *If ϵ is a homomorphism, let $K(\sqrt{\Delta})/K$ be the fixed field of the kernel of ϵ. For each $v \in \Sigma$ and quadratic character $\chi \in \mathcal{C}(K_v)$ write L_χ/K_v for the extension cut out by χ and define*

$$\omega_v(\chi) := \chi(\Delta)(-1)^{\dim_{\mathbb{F}_2} A(L_\chi)/N_{L_\chi/K_v} A(L_\chi)}.$$

Finally, define

$$\delta_v := \frac{1}{|\mathcal{C}(K_v)|} \sum_{\chi \in \mathcal{C}(K_v)} \omega(\chi) \quad \text{and} \quad \delta := \prod_{v \in \Sigma} \delta_v.$$

Then for all sufficiently large X,

$$\frac{|\{\chi \in \mathcal{C}(K, X) : \dim_{\mathbb{F}_2} \mathrm{Sel}_2(A^\chi/K) \text{ is even}\}|}{|\mathcal{C}(K, X)|} = \frac{1 + (-1)^{\dim_{\mathbb{F}_2} \mathrm{Sel}_2(A/K)} \cdot \delta}{2}.$$

Proof. Combine Proposition 10.11, Theorem 7.4 and Lemma 10.8. □

Remark 10.14. Lemma 10.9(ii) enables one to evaluate the local terms δ_v for archimedean places. For nonarchimedean places of odd residue characteristic, the dimension of the cokernel of the norm map may be expressed in terms of Tamagawa numbers, see [Morgan 2015, Lemma 2.5].

In the following examples we examine when ϵ is (or is not) a homomorphism for certain families of abelian varieties.

Example 10.15 (generic 2-torsion). For any principally polarized abelian variety A/K of dimension g, $\mathrm{Gal}(K(A[2])/K)$ is a subgroup of the symplectic group $\mathrm{Sp}_{2g}(\mathbb{F}_2)$. As in Remark 7.5, if $g \geq 2$ and $\mathrm{Gal}(K(A[2])/K) \cong \mathrm{Sp}_{2g}(\mathbb{F}_2)$ then ϵ is not a homomorphism.

Example 10.16 (elliptic curves). Suppose that A/K is an elliptic curve, say given by a Weierstrass equation of the form $y^2 = f(x)$ for some monic (separable) cubic polynomial $f(x)$. Then $\mathrm{Gal}(K(A[2])/K) = \mathrm{Gal}(f)$ is the Galois group of the splitting field of $f(x)$ and as such may be viewed as a subgroup of the symmetric group S_3. One readily checks that the map $\sigma \mapsto (-1)^{\dim_{\mathbb{F}_2} A[2]^\sigma}$ is the sign homomorphism. Thus ϵ is always a homomorphism and we may take Δ to be the discriminant of the elliptic curve. Thus Theorem 10.13 recovers [Klagsbrun et al. 2013, Theorem A]. See Proposition 7.9 of that work for a table computing the local terms δ_v as a function of the reduction of the elliptic curve.

Example 10.17 (hyperelliptic curves). Let C/K be a hyperelliptic curve of genus $g \geq 2$, say given by a Weierstrass equation $y^2 = f(x)$ for a (separable, not necessarily monic) polynomial $f(x)$ with $\deg(f) \in \{2g+1, 2g+2\}$. Take A/K to be the Jacobian of C so that A/K is a principally polarized abelian variety of dimension g. Then again $\mathrm{Gal}(K(A[2])/K) = \mathrm{Gal}(f)$ which we view as a subgroup of the symmetric group $S_{\deg(f)}$. Write $\mathrm{sgn} : S_n \to \{\pm 1\}$ for the sign homomorphism and fix $\sigma \in \mathrm{Gal}(f)$ with cycle type $(d_1 \cdots d_s)$. Then we have

$$\epsilon(\sigma) = \begin{cases} -\mathrm{sgn}(\sigma) & \text{all } d_i \text{ even and } \deg(f) \pmod 2, \\ \mathrm{sgn}(\sigma) & \text{else.} \end{cases} \tag{10.18}$$

Indeed, this follows from [Cornelissen 2001, Theorem 1.4] (whilst [loc. cit.] is stated for hyperelliptic curves over finite fields of odd residue characteristic, the proof yields the above statement for all fields of characteristic not 2; note also the erratum [Cornelissen 2005]).

Suppose now that either g is odd or $\deg(f)$ is odd. Then by (10.18) ϵ is always a homomorphism and again we may take Δ to be the discriminant of the hyperelliptic curve C. In particular, the case $\deg(f)$ odd recovers [Yu 2016, Theorem 1].

Now suppose that both g and $\deg(f)$ are even, or equivalently $\deg(f) \equiv 2 \pmod 4$. Suppose further that either $\mathrm{Gal}(f) \cong S_{2g+2}$ or $\mathrm{Gal}(f) \cong A_{2g+2}$. Then by (10.18) we see that ϵ is not a homomorphism (indeed, the only nontrivial homomorphism from S_{2g+2} to $\{\pm 1\}$ is sgn yet (10.18) shows that ϵ is nontrivial when restricted to A_{2g+2}).

Example 10.19 (abelian varieties with principal polarization induced by a rational symmetric line bundle). Suppose that $(A/K, \lambda)$ is a principally polarized abelian variety and that the polarization λ is induced by a rational (i.e., G_K-invariant) symmetric line bundle $\mathcal{L}$. Then the associated quadratic refinement $q_{\mathcal{L}}$ of the Weil-pairing $(\cdot,\cdot)_\lambda$ on $A[2]$ (as in Definition 4.3) is G_K-invariant also, whence $\mathrm{Gal}(K(A[2])/K)$ acts on $A[2]$ through the orthogonal group $O(q_{\mathcal{L}})$. Then ϵ is the Dickson homomorphism $d_{q_{\mathcal{L}}}$ (Proposition 3.5). We remark that this case includes both elliptic curves and Jacobians of hyperelliptic curves of either odd degree or odd genus, see [Poonen and Rains 2011, Proposition 3.11].

10D. ***Main theorems for 2^∞-Selmer ranks.*** We now incorporate the results of Section 5 to move from 2-Selmer ranks to 2^∞-Selmer ranks.

Theorem 10.20. *Let K be a number field and $(A/K, \lambda)$ a principally polarized abelian variety. Let Σ be the set consisting of all archimedean places of K, all places of bad reduction for A, and all places dividing 2, and let L/K be a quadratic extension with associated quadratic character χ. Then*

$$\mathrm{rk}_2(A/L) \equiv \sum_{\substack{v\in\Sigma \\ v \text{ nonsplit in } L/K}} (2\,\mathrm{inv}_v\, \mathfrak{g}(A/K_v, \lambda_v, \chi_v) + \dim_{\mathbb{F}_2} A(K_v)/N_{L_w/K_v} A(L_w)) \pmod 2$$

where the local terms [2] $\mathfrak{g}(A/K_v, \lambda_v, \chi_v) \in \mathrm{Br}(K_v)[2]$ *are given in Definition 5.15, and w denotes any place of L extending v.*

Proof. First note that $\mathrm{rk}_2(A/L) = \mathrm{rk}_2(A/K) + \mathrm{rk}_2(A^\chi/K)$. Moreover, we have

$$\dim_{\mathbb{F}_2} \mathrm{Sel}_2(A/K) = \mathrm{rk}_2(A/K) + \dim_{\mathbb{F}_2} A(K)[2] + \dim_{\mathbb{F}_2} Ш_{\mathrm{nd}}(A/K)[2]$$

and the analogous equality for A^χ/K. Noting that $\dim_{\mathbb{F}_2} A(K)[2] = \dim_{\mathbb{F}_2} A^\chi(K)[2]$ the above observations combine to give

$$\begin{aligned}&\mathrm{rk}_2(A/L)\\ &\quad\equiv \dim_{\mathbb{F}_2} \mathrm{Sel}_2(A/K) + \dim_{\mathbb{F}_2} \mathrm{Sel}_2(A^\chi/K) + \dim_{\mathbb{F}_2} Ш_{\mathrm{nd}}(A/K)[2] + \dim_{\mathbb{F}_2} Ш_{\mathrm{nd}}(A^\chi/K)[2] \pmod 2.\end{aligned}$$

Combining Theorem 10.12 with Theorem 5.20 then gives

$$\mathrm{rk}_2(A/L) \equiv \sum_{v\in M_K} (2\,\mathrm{inv}_v\, \mathfrak{g}(A/K_v, \lambda_v, \chi_v) + \dim_{\mathbb{F}_2} A(K_v)/N_{L_w/K_v} A(L_w)) \pmod 2.$$

Finally, combining Proposition 5.16 with Lemma 10.9 shows that

$$2\,\mathrm{inv}_v\, \mathfrak{g}(A/K_v, \lambda_v, \chi_v) + \dim_{\mathbb{F}_2} A(K_v)/N_{L_w/K_v} A(L_w) \equiv 0 \pmod 2$$

for each place $v \notin \Sigma$, and similarly for each place $v \in \Sigma$ which split in L/K. □

We now prove Theorem 1.2, after first defining the local terms appearing in the statement.

Definition 10.21. Let K be a number field, $(A/K, \lambda)$ a principally polarized abelian variety, and let Σ denote the set consisting of all archimedean places of K, all places of bad reduction for A, and all places dividing 2.

For each $v \in \Sigma$ and $\chi \in \mathcal{C}(K_v)$ define

$$\Omega_v(\chi) := (-1)^{2\,\mathrm{inv}_v\, \mathfrak{g}(A/K_v, \lambda_v, \chi) + \dim_{\mathbb{F}_2} A(K_v)/N_{L_\chi/K_v} A(L_\chi)}$$

[2] Here and in Definition 10.21 we think of $\mathrm{inv}_v\, \mathfrak{g}(A/K_v, \lambda_v, \chi_v)$ as being equal to 0 or $\frac{1}{2}$ (as opposed to the class of this in $\mathbb{Q}/\mathbb{Z}$) so that $2\,\mathrm{inv}_v\, \mathfrak{g}(A/K_v, \lambda_v, \chi_v)$ is either 0 or 1 accordingly.

where here L_χ is the extension of K_v cut out by χ. Further, we define (for each $v \in \Sigma$)

$$\kappa_v = \frac{1}{|\mathcal{C}(K_v)|} \sum_{\chi \in \mathcal{C}(K_v)} \Omega_v(\chi) \quad \text{and} \quad \kappa = \prod_{v \in \Sigma} \kappa_v.$$

Remark 10.22. If v is archimedean then by Theorem 10.20 we have

$$\Omega_v(\chi_v) = \begin{cases} 1 & \chi_v \text{ trivial,} \\ (-1)^{\dim A} & \text{else.} \end{cases}$$

In particular, if v is a real place and $\dim A$ is odd then $\kappa_v = 0$ (hence also $\kappa = 0$), whilst if v is complex or $\dim A$ is even, we have $\kappa_v = 1$.

Theorem 10.23. *Let A/K be a principally polarized abelian variety. Then for all sufficiently large $X > 0$,*

$$\frac{|\{\chi \in \mathcal{C}(K, X) : \mathrm{rk}_2(A^\chi/K) \text{ is even}\}|}{|\mathcal{C}(K, X)|} = \frac{1 + (-1)^{\mathrm{rk}_2(A/K)} \cdot \kappa}{2}.$$

Proof. As noted previously, for any $\chi \in \mathcal{C}(K)$ corresponding to the quadratic extension L/K, we have

$$\mathrm{rk}_2(A/L) = \mathrm{rk}_2(A/K) + \mathrm{rk}_2(A^\chi/K).$$

Thus for each $\chi \in \mathcal{C}(K)$, Theorem 10.20 gives

$$(-1)^{\mathrm{rk}_2(A^\chi/K)} = (-1)^{\mathrm{rk}_2(A/K)} \prod_{v \in \Sigma} \Omega(\chi_v)$$

with $\Omega(\chi_v) \in \{\pm 1\}$ depending only on the restriction of χ to K_v. The argument is now identical to that in the proof of Theorem 7.10. As is the case there, "sufficiently large $X > 0$" means that we require only that X is large enough that the restriction homomorphism from $\mathcal{C}(K, X)$ to $\prod_{v \in \Sigma} \mathcal{C}(K_v)$ is surjective. □

The following example shows that the proportion of twists having even 2-Selmer rank can differ from the proportion having even 2^∞-Selmer rank.

Example 10.24. Consider the genus 2 hyperelliptic curve $C : y^2 = x^6 + x^4 + x + 3$ over $\mathbb{Q}$. The polynomial $f(x) = x^6 + x^4 + x + 3$ has Galois group S_6. By Theorem 10.13 (see also Example 10.17) the 2-Selmer ranks are distributed half-and-half amongst even/odd in the quadratic twist family of the Jacobian J/K of C.

On the other hand, we claim that $\kappa = \frac{3}{16}$ so that $\frac{19}{32}$ of the twists of J have even 2^∞-Selmer rank whilst $\frac{13}{32}$ have odd 2^∞-Selmer rank. The discriminant of $f(x)$ is $-5 \cdot 2670719$, so J/K has good reduction away from 2, 5 and 2670719. Thus we have $\Sigma = \{2, 5, 2670719, \infty\}$. Using the computer algebra package MAGMA [Bosma et al. 1997], one computes that $\mathrm{rk}_2(J/K)$ is odd. By Remark 10.22, $\kappa_\infty = 1$. To compute κ_2, κ_5 and $\kappa_{2670719}$, one may use the following trick. By Theorem 10.20 and the above discussion, for a quadratic character χ of $\mathbb{Q}$ corresponding to the extension $L/\mathbb{Q}$, one has

$$(-1)^{\mathrm{rk}_2(J^\chi/\mathbb{Q})} = - \prod_{\substack{v \in \{2,5,2670719\} \\ v \text{ nonsplit in } L}} \Omega_v(\chi_v). \tag{10.25}$$

Now for $0 \neq n \in \mathbb{Z}$, the quadratic twist of J by $\mathbb{Q}(\sqrt{n})/\mathbb{Q}$ is the Jacobian of the hyperelliptic curve $y^2 = nf(x)$. Thus one may use MAGMA to compute $(-1)^{\mathrm{rk}_2(J^\chi/\mathbb{Q})}$ for various (finitely many) quadratic characters χ, from which one may then determine all the $\Omega_v(\chi_v)$ by (10.25). Upon doing this one obtains $\kappa_2 = \frac{3}{4}$, $\kappa_5 = -\frac{1}{2}$ and $\kappa_{2670719} = \frac{1}{2}$ and the claim follows.

Remark 10.26. Since by Theorem 10.20 the parity of $\mathrm{rk}_2(A^\chi/K)$ depends only on the restriction of χ to the archimedean places, the places of bad reduction for A, and the places over 2, it follows from Theorem 9.4(i) (along with Proposition 10.11) that when ϵ fails to be a homomorphism we in fact have

$$\frac{|\{\chi \in \mathcal{C}(K, X) : \dim_{\mathbb{F}_2} \mathrm{Sel}_2(A^\chi/K) \text{ is even and } \mathrm{rk}_2(A^\chi/K) \text{ is even}\}|}{|\{\chi \in \mathcal{C}(K, X) : \mathrm{rk}_2(A^\chi/K) \text{ is even}\}|} = \frac{1}{2}$$

for all sufficiently large X (assuming the denominator is nonzero) and that the same holds when we condition on $\mathrm{rk}_2(A^\chi/K)$ being odd also. Thus when ϵ fails to be a homomorphism the parities of Selmer ranks and the parities of 2-infinity Selmer ranks behave "independently".

10E. ***The proportion of twists having nonsquare Shafarevich–Tate group.*** We now prove an analogue of Theorem 1.1 for $\dim_{\mathbb{F}_2} Ш_{\mathrm{nd}}(A/K)[2]$ rather than for $\dim_{\mathbb{F}_2} \mathrm{Sel}_2(A/K)$. Since the Shafarevich–Tate group of a principally polarized abelian variety, if finite, has square order if and only if $\dim_{\mathbb{F}_2} Ш_{\mathrm{nd}}(A/K)[2]$ is even (see e.g., [Poonen and Stoll 1999, Theorem 8]), this may be viewed as quantifying the failure of the Shafarevich–Tate group to have square order in quadratic twist families. The proof of the theorem is identical to its analogue for 2-Selmer ranks, so we only sketch the proof.

Theorem 10.27. *Let K be a number field, $(A/K, \lambda)$ a principally polarized abelian variety, and Σ the set consisting of all archimedean places of K, all places of bad reduction for A, and all places dividing 2. Define $\epsilon : \mathrm{Gal}(K(A[2])/K) \to \{\pm 1\}$ by $\sigma \mapsto (-1)^{\dim_{\mathbb{F}_2} A[2]^\sigma}$.*

(i) *If ϵ fails to be a homomorphism then for all sufficiently large X*

$$\frac{|\{\chi \in \mathcal{C}(K, X) : \dim_{\mathbb{F}_2} Ш_{\mathrm{nd}}(A^\chi/K)[2] \textit{ is even}\}|}{|\mathcal{C}(K, X)|} = \frac{1}{2}.$$

(ii) *If ϵ is a homomorphism, let $K(\sqrt{\Delta})/K$ be the fixed field of the kernel of ϵ. For each $v \in \Sigma$ and quadratic character $\chi \in \mathcal{C}(K_v)$ write L_χ/K_v for the extension cut out by χ and define*

$$\Upsilon_v(\chi) := \chi(\Delta)(-1)^{2\,\mathrm{inv}_v\, \mathfrak{g}(A/K_v, \lambda_v, \chi_v)}.$$

Finally, define

$$\rho_v := \frac{1}{|\mathcal{C}(K_v)|} \sum_{\chi \in \mathcal{C}(K_v)} \Upsilon(\chi) \quad \textit{and} \quad \rho := \prod_{v \in \Sigma} \rho_v.$$

Then for all sufficiently large X,

$$\frac{|\{\chi \in \mathcal{C}(K, X) : \dim_{\mathbb{F}_2} Ш_{\mathrm{nd}}(A^\chi/K)[2] \textit{ is even}\}|}{|\mathcal{C}(K, X)|} = \frac{1 + (-1)^{\dim_{\mathbb{F}_2} Ш_{\mathrm{nd}}(A/K)[2]} \cdot \rho}{2}.$$

Proof. Fix a quadratic character χ of K. As in Definition 9.1, set

$$w(\chi) := \prod_{v\notin\Sigma,\,\chi_v \text{ ram}} (-1)^{\dim_{\mathbb{F}_p} A(K_v)[2]} = \prod_{v\notin\Sigma,\,\chi_v \text{ ram}} \epsilon(\mathrm{Frob}_v).$$

Combining Theorem 5.20 with Proposition 5.16 we obtain

$$(-1)^{\dim_{\mathbb{F}_2} Ш_{\mathrm{nd}}(A^\chi/K)[2]} = w(\chi)(-1)^{\dim_{\mathbb{F}_2} Ш_{\mathrm{nd}}(A/K)[2]} \prod_{v\in\Sigma} (-1)^{2\,\mathrm{inv}_v\, \mathfrak{g}(A/K_v,\lambda_v,\chi_v)}. \tag{10.28}$$

If ϵ is a homomorphism then, as in the proof of Lemma 7.7, we have $w(\chi) = \prod_{v\in\Sigma} \chi_v(\Delta)$, whence

$$(-1)^{\dim_{\mathbb{F}_2} Ш_{\mathrm{nd}}(A^\chi/K)[2]} = (-1)^{\dim_{\mathbb{F}_2} Ш_{\mathrm{nd}}(A/K)[2]} \prod_{v\in\Sigma} \Upsilon_v(\chi)$$

and the same argument as in the proof of Theorem 7.10 gives the result.

On the other hand, suppose that ϵ is a homomorphism and enlarge Σ if necessary so that Assumption 8.8 holds, noting that (10.28) still remains true. The result now follows from Theorem 9.4 (cf. proof of Theorem 9.5). □

10F. ***The joint distribution of parities of 2-Selmer ranks and 2-infinity Selmer ranks.*** By combining Theorem 10.27 with Theorems 10.13 and 10.23 we are able to push Remark 10.26 further to determine the "joint distribution" of parities of 2-Selmer ranks and 2-infinity Selmer ranks.

Corollary 10.29 (of Theorem 10.27). *Let K be a number field, A/K a principally polarized abelian variety, and $\epsilon : \mathrm{Gal}(K(A[2])/K) \to \{\pm 1\}$ the map $\sigma \mapsto (-1)^{\dim_{\mathbb{F}_2} A[2]^\sigma}$. Let the constants δ, κ and ρ be as in Theorem 10.13, Definition 10.21 and Theorem 10.27 respectively. Then for $m, n \in \{0,1\}$ we have, for all sufficiently large X,*

$$\big|\{\chi \in \mathcal{C}(K,X) : \mathrm{rk}_2(A^\chi/K) \equiv m \pmod 2, \dim_{\mathbb{F}_2} \mathrm{Sel}_2(A^\chi/K) \equiv n \pmod 2\}\big| / |\mathcal{C}(K,X)| = \tfrac{1}{4} + (-1)^m a_1 + (-1)^n a_2 + (-1)^{m+n} a_3$$

where

$$a_1 = \tfrac{1}{4}(-1)^{\mathrm{rk}_2(A/K)}\kappa,$$

and $a_2 = a_3 = 0$ if ϵ fails to be a homomorphism whilst

$$a_2 = \tfrac{1}{4}(-1)^{\dim_{\mathbb{F}_2} \mathrm{Sel}_2(A/K)}\delta \quad \textit{and} \quad a_3 = \tfrac{1}{4}(-1)^{\dim_{\mathbb{F}_2} \mathrm{Sel}_2(A/K)+\mathrm{rk}_2(A/K)}\rho$$

otherwise.

Proof. Follows from Theorems 10.27, 10.13 and 10.23 upon noting that, for any $\chi \in \mathcal{C}(K)$, we have

$$\dim_{\mathbb{F}_2} \mathrm{Sel}_2(A^\chi/K) = \mathrm{rk}_2(A^\chi/K) + \dim_{\mathbb{F}_2} A(K)[2] + \dim_{\mathbb{F}_2} Ш_{\mathrm{nd}}(A^\chi/K)[2].$$ □

11. Twisting data for abelian varieties ($p > 2$)

As in the previous section, let K be a number field and $(A/K, \lambda)$ a principally polarized abelian variety. This time we take p to be an odd prime and $T = A[p]$. As with $A[2]$ in the previous section, we endow T with a canonical global metabolic structure and twisting data so that the resulting Selmer groups have a classical interpretation. For elliptic curves this is done by Klagsbrun, Mazur and Rubin [2013, §5]. This time the case of an arbitrary principally polarized abelian variety is almost identical to that of [loc. cit.], though to fix notation we repeat the relevant material.

11A. ***The global metabolic structure on*** $A[p]$**.** As with the case $p = 2$, the polarization λ along with the Weil-pairing

$$(\cdot\,,\cdot)_{e_p} : A[p] \times A^\vee[p]$$

provides the desired (nondegenerate, alternating, G_K-equivariant) bilinear pairing

$$(\cdot\,,\cdot)_\lambda : T \times T \to \boldsymbol{\mu}_p$$

(defined by setting $(x, y)_\lambda = (x, \lambda(y))_{e_p}$ for $x, y \in T$). We take Σ to be a finite set of places of K containing all archimedean places, all primes over p, and all primes at which A has bad reduction. Then T is unramified outside Σ.

Since p is odd, the quadratic forms $q_v = \frac{1}{2}\langle\cdot\,,\cdot\rangle_v$ (here v a place of K and $\langle\cdot\,,\cdot\rangle_v$ denotes the local Tate pairing associated to $(\cdot\,,\cdot)_\lambda$) are Tate quadratic forms which endow T with a global metabolic structure $\boldsymbol{q}$ (cf. Section 6C).

11B. ***Twisting data associated to*** $A[p]$**.** Here we associate canonical twisting data to $(A[p], \Sigma, \boldsymbol{q})$.

Definition 11.1. Let $\chi \in \mathcal{C}(K)$ be nontrivial and let L denote the associated cyclic p-extension $L = \overline{K}^{\ker(\chi)}$ of K. We write A^χ for the abelian variety denoted A_L in [Mazur et al. 2007, Definition 5.1], so that A^χ/K is an abelian variety of dimension $(p-1)\dim A$ which may be defined as the kernel of the "norm" homomorphism $\mathrm{Res}_{L/K} A \to A$ (here $\mathrm{Res}_{L/K} A$ denotes the restriction of scalars of A from L to K).

By [Mazur et al. 2007, Theorem 5.5(iv)], χ induces an inclusion of $\mathbb{Z}[\boldsymbol{\mu}_p]$ into $\mathrm{End}_K(A^\chi)$. Moreover, by Theorem 2.2(iii) of [loc. cit.] we have a canonical isomorphism $\psi : A[p] \xrightarrow{\sim} A^\chi[\mathfrak{p}]$ where $\mathfrak{p}$ denotes the unique prime of $\mathbb{Z}[\boldsymbol{\mu}_p]$ lying over p.

If $\mathbb{1}_{K_v} \neq \chi \in \mathcal{C}(K_v)$ for some place of K then we define A^χ/K_v similarly.

Remark 11.2. Fix $\chi \in \mathcal{C}(K)$ nontrivial, and let π be a generator of the prime $\mathfrak{p}$ of $\mathbb{Z}[\boldsymbol{\mu}_p]$ lying over p. View π inside $\mathrm{End}_K(A^\chi)$ as above. Then π is an isogeny and we have an associated π-Selmer group

$$\mathrm{Sel}_\pi(A^\chi/K) = \{\mathfrak{a} \in H^1(K, A^\chi[\mathfrak{p}]) : \mathfrak{a}_v \in \mathrm{im}(\delta_v)\ \forall v \in M_K\},$$

where here for each place v of K, $\delta_v : A^\chi(K_v)/\pi A^\chi(K_v) \hookrightarrow H^1(K_v, A^\chi[\mathfrak{p}])$ is the connecting homomorphism associated to the multiplication-by-π Kummer sequence for A^χ/K_v.

One checks that $\mathrm{Sel}_\pi(A^\chi/K)$ does not depend on the choice of generator π for $\mathfrak{p}$.

We now define the twisting data.

Definition 11.3. Let v be a place of K and $\chi \in \mathcal{C}(K_v)$. Define $\alpha_v(\chi) \subseteq H^1(K_v, A[p])$ as follows:

(i) If χ is trivial, define $\alpha_v(\chi)$ to be the image of $A(K)/pA(K)$ under the connecting homomorphism associated to the multiplication-by-p Kummer sequence for A/K_v,

(ii) If χ is nontrivial, let π be a generator of the prime $\mathfrak{p}$ of $\mathbb{Z}[\boldsymbol{\mu}_p]$ lying over p. Then we define $\alpha_v(\chi)$ to be the image of $A^\chi(K_v)/\pi A^\chi(K_v)$ under the composition

$$A^\chi(K_v)/\pi A^\chi(K_v) \xrightarrow{\delta_v} H^1(K_v, A^\chi[\mathfrak{p}]) \xrightarrow{\sim} H^1(K_v, A[p]),$$

where the rightmost map is induced by the isomorphism $\psi : A[p] \xrightarrow{\sim} A[\mathfrak{p}]$ of Definition 11.1. One sees easily that $\alpha_v(\chi)$ does not depend on the choice of π, and depends only on the extension cut out by χ.

As usual, for v a place of K and $\chi_1, \chi_2 \in \mathcal{C}(K_v)$, write

$$h_v(\chi_1, \chi_2) = \dim_{\mathbb{F}_p}(\alpha_v(\chi_1)/(\alpha_v(\chi_1) \cap \alpha_v(\chi_2))).$$

As in the case $p = 2$, we have.

Lemma 11.4. *Let v be a place of K, $\chi \in \mathcal{C}(K_v)$ and L_χ the extension of K_v cut out by χ. Then*

$$h_v(\mathbb{1}_{K_v}, \chi) = \dim_{\mathbb{F}_p} A(K_v)/N_{L_\chi/K_v} A(L_\chi)$$

where $N_{L_\chi/K_v} : A(L_\chi) \to A(K_v)$ is the norm map.

Moreover, if $v \nmid p$ is a nonarchimedean place of K at which A has good reduction then:

(i) *If χ is unramified, we have*

$$h_v(\mathbb{1}_{K_v}, \chi) = \dim_{\mathbb{F}_p} A(K_v)/N_{L_\chi/K_v} A(L_\chi) = 0.$$

(ii) *If χ is ramified, we have*

$$h_v(\mathbb{1}_{K_v}, \chi) = \dim_{\mathbb{F}_p} A(K_v)/N_{L_\chi/K_v} A(L_\chi) = \dim_{\mathbb{F}_p} A(K_v)[p].$$

Proof. As in the case $p = 2$ the first claim is shown for elliptic curves in [Mazur and Rubin 2007, Proposition 5.2] and the argument is identical. The evaluation of the cokernel of the local norm map is [Mazur 1972, Corollary 4.4] for χ unramified, and for χ ramified the case where A is an elliptic curve is [Mazur and Rubin 2007, Lemma 5.5(ii)] and the same argument works in general. □

Proposition 11.5. *The maps $\boldsymbol{\alpha} = (\alpha_v)_v$ define twisting data for $T = (A[p], \boldsymbol{q}, \Sigma)$ and the associated Selmer groups* $\mathrm{Sel}(A[p], \chi)$ *satisfy*

$$\mathrm{Sel}(A[p], \chi) \cong \mathrm{Sel}_\pi(A^\chi/K).$$

Proof. We first claim that for each place v of K and $\chi \in \mathcal{C}(K_v)$, we have $\alpha_v(\chi) \in \mathcal{H}(q_v)$ (i.e., $\alpha_v(\chi)$ is Lagrangian). That is (since p is odd), that $\alpha_v(\chi) \subseteq H^1(K_v, A[p])$ is its own orthogonal complement under the Tate pairing. For χ trivial, that (the image in $H^1(K_v, A[p])$ of) $A(K_v)/pA(K_v)$ is its own orthogonal complement is a well known consequence of Tate local duality, see e.g., [Milne 2006, I.3.4]. For χ nontrivial this is shown for A an elliptic curve in [Mazur and Rubin 2007, Proposition A.7] and the argument for a general principally polarized abelian variety is identical (with the Weil pairing associated to the principal polarization λ providing the pairing on the p-adic Tate-module $T_p(A)$ required for Definition A.5 of [loc. cit.]). We remark that in the above, unlike the case $p = 2$, the twist A^χ need not possess a principal polarization (see [Howe 2001, Theorem 1.1]) so one cannot deduce the result by just applying Tate duality to A^χ/K_v, as one does not have an appropriate Weil-pairing on $A[\mathfrak{p}]$.

To show that $\boldsymbol{\alpha}$ defines twisting data, it remains to show that for each place $v \notin \Sigma$ with $\boldsymbol{\mu}_p \subseteq K_v$, we have $\alpha_v(\chi) \in \mathcal{H}_{\mathrm{ram}}(q_v)$. That is, that $\alpha_v(\chi) \cap H^1_{\mathrm{ur}}(K_v, A[p]) = 0$. Again, the argument is the same as in the case $p = 2$. Indeed, for such places we have $\alpha_v(\mathbb{1}_{K_v}) = H^1_{\mathrm{ur}}(K_v, A[p])$ (again, see e.g., [Poonen and Rains 2012, Proposition 4.12] and the preceding remark) and we conclude by Lemma 11.4(ii).

The isomorphism $\mathrm{Sel}(A[p], \chi) \cong \mathrm{Sel}_\pi(A^\chi/K)$ is also proven identically to the case $p = 2$ by comparing the local conditions defining the two Selmer groups. □

Corollary 11.6 (Theorem 1.5). *Let p be an odd prime, K a number field, A/K a principally polarized abelian variety, and Σ the set consisting of all archimedean places of K, all places of bad reduction for A, and all places dividing p. Define $\epsilon : \mathrm{Gal}(K(A[p])/K) \to \{\pm 1\}$ by $\sigma \mapsto (-1)^{\dim_{\mathbb{F}_p} A[p]^\sigma}$.*

(i) *If ϵ is nontrivial when restricted to $\mathrm{Gal}(K(A[p])/K(\boldsymbol{\mu}_p))$ then*

$$\lim_{X\to\infty} \frac{|\{\chi \in \mathcal{C}(K, X) : \dim_{\mathbb{F}_p} \mathrm{Sel}_\pi(A^\chi/K) \text{ is even}\}|}{|\mathcal{C}(K, X)|} = \frac{1}{2}.$$

(ii) *Suppose ϵ is trivial when restricted to $\mathrm{Gal}(K(A[p])/K(\boldsymbol{\mu}_p))$. For each $v \in \Sigma$ and character $\chi \in \mathcal{C}(K_v)$, write L_χ/K_v for the extension cut out by χ and define*

$$\omega_v(\chi) := (-1)^{\dim_{\mathbb{F}_p} A(L_\chi)/N_{L_\chi/K_v} A(L_\chi)}.$$

Finally, define

$$\delta_v := \frac{1}{|\mathcal{C}(K_v)|} \sum_{\chi \in \mathcal{C}(K_v)} \omega(\chi) \quad \text{and} \quad \delta := \prod_{v\in\Sigma} \delta_v.$$

Then for all sufficiently large X,

$$\frac{|\{\chi \in \mathcal{C}(K, X) : \dim_{\mathbb{F}_p} \mathrm{Sel}_\pi(A^\chi/K) \text{ is even}\}|}{|\mathcal{C}(K, X)|} = \frac{1 + (-1)^{\dim_{\mathbb{F}_p} \mathrm{Sel}_p(A/K)} \cdot \delta}{2}.$$

Proof. Combine Theorem 7.4 with Proposition 11.5 and Lemma 11.4. □

Remark 11.7. As observed by Klagsbrun, Mazur and Rubin [2013, immediately before the statement of Theorem 8.2], as each $|\mathcal{C}(K_v)|$ has odd size we cannot have $\delta = 0$ in case (ii) above.

Remark 11.8. As in Remark 7.5, a sufficient condition to ensure that ϵ is nontrivial when restricted to $\mathrm{Gal}(K(A[p])/K(\boldsymbol{\mu}_p))$ is that $\mathrm{Gal}(K(A[p])/K)$ (viewed as a subgroup of $\mathrm{GSp}_{2g}(\mathbb{F}_p)$ for $g = \dim A$) contains a symplectic transvection. In particular, if the Galois action on $A[p]$ is as large as possible, so that $\mathrm{Gal}(K(A[p])/K) \cong \mathrm{GSp}_{2g}(\mathbb{F}_p)$, then case (i) of Corollary 11.6 applies. It is also known that $\mathrm{Gal}(K(A[p])/K)$ contains a transvection if there is a place v of K, not dividing p, such that A has semistable reduction of toric dimension 1 at v, and such that the order of the Néron component group of A/K_v is coprime to p (see [Le Duff 1998, Proposition 1.3]).

Acknowledgements

We thank Kęstutis Česnavičius for many useful conversations and comments, and for correspondence regarding the material in Section 4. We thank Tim Dokchitser, Vladimir Dokchitser, Céline Maistret and Jeremy Rickard for helpful conversations, and the anonymous referee for pointing out Corollary 10.29. We would like to thank the University of Warwick and King's College London where parts of this research were carried out. This research is supported by EPSRC grant EP/M016846.

References

[Atiyah and Wall 1967] M. F. Atiyah and C. T. C. Wall, "Cohomology of groups", pp. 94–115 in *Algebraic number theory* (Brighton, 1965), edited by J. W. S. Cassels and A. Fröhlich, Thompson, Washington, D.C., 1967. MR

[Bosma et al. 1997] W. Bosma, J. Cannon, and C. Playoust, "The Magma algebra system, I: The user language", *J. Symbolic Comput.* **24**:3-4 (1997), 235–265. MR Zbl

[Cornelissen 2001] G. Cornelissen, "Two-torsion in the Jacobian of hyperelliptic curves over finite fields", *Arch. Math. (Basel)* **77**:3 (2001), 241–246. MR Zbl

[Cornelissen 2005] G. Cornelissen, "Erratum to: "Two-torsion in the Jacobian of hyperelliptic curves over finite fields" [Arch. Math. (Basel) **77** (2001), no. 3, 241–246; MR1865865]", *Arch. Math. (Basel)* **85**:6 (2005), loose erratum. MR Zbl

[Dye 1977] R. H. Dye, "A geometric characterization of the special orthogonal groups and the Dickson invariant", *J. London Math. Soc.* (2) **15**:3 (1977), 472–476. MR Zbl

[Flach 1990] M. Flach, "A generalisation of the Cassels–Tate pairing", *J. Reine Angew. Math.* **412** (1990), 113–127. MR Zbl

[Howe 2001] E. W. Howe, "Isogeny classes of abelian varieties with no principal polarizations", pp. 203–216 in *Moduli of abelian varieties* (Texel Island, 1999), edited by C. Faber et al., Progr. Math. **195**, Birkhäuser, Basel, 2001. MR Zbl

[Klagsbrun et al. 2013] Z. Klagsbrun, B. Mazur, and K. Rubin, "Disparity in Selmer ranks of quadratic twists of elliptic curves", *Ann. of Math.* (2) **178**:1 (2013), 287–320. MR Zbl

[Klagsbrun et al. 2014] Z. Klagsbrun, B. Mazur, and K. Rubin, "A Markov model for Selmer ranks in families of twists", *Compos. Math.* **150**:7 (2014), 1077–1106. MR Zbl

[Kramer 1981] K. Kramer, "Arithmetic of elliptic curves upon quadratic extension", *Trans. Amer. Math. Soc.* **264**:1 (1981), 121–135. MR Zbl

[Le Duff 1998] P. Le Duff, "Représentations galoisiennes associées aux points d'ordre l des jacobiennes de certaines courbes de genre 2", *Bull. Soc. Math. France* **126**:4 (1998), 507–524. MR Zbl

[Mazur 1972] B. Mazur, "Rational points of abelian varieties with values in towers of number fields", *Invent. Math.* **18** (1972), 183–266. MR Zbl

[Mazur and Rubin 2007] B. Mazur and K. Rubin, "Finding large Selmer rank via an arithmetic theory of local constants", *Ann. of Math.* (2) **166**:2 (2007), 579–612. MR Zbl

[Mazur et al. 2007] B. Mazur, K. Rubin, and A. Silverberg, "Twisting commutative algebraic groups", *J. Algebra* **314**:1 (2007), 419–438. MR Zbl

[Milne 1986] J. S. Milne, "Abelian varieties", pp. 103–150 in *Arithmetic geometry* (Storrs, Conn., 1984), edited by G. Cornell and J. H. Silverman, Springer, 1986. MR Zbl

[Milne 2006] J. S. Milne, *Arithmetic duality theorems*, 2nd ed., BookSurge, LLC, Charleston, SC, 2006. MR Zbl

[Morgan 2015] A. Morgan, "2-Selmer parity for hyperelliptic curves in quadratic extensions", 2015. arXiv

[Mumford 1966] D. Mumford, "On the equations defining abelian varieties, I", *Invent. Math.* **1** (1966), 287–354. MR Zbl

[Neukirch et al. 2008] J. Neukirch, A. Schmidt, and K. Wingberg, *Cohomology of number fields*, 2nd ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences] **323**, Springer, 2008. MR Zbl

[Pollatsek 1971] H. Pollatsek, "First cohomology groups of some linear groups over fields of characteristic two", *Illinois J. Math.* **15** (1971), 393–417. MR Zbl

[Poonen and Rains 2011] B. Poonen and E. Rains, "Self cup products and the theta characteristic torsor", *Math. Res. Lett.* **18**:6 (2011), 1305–1318. MR Zbl

[Poonen and Rains 2012] B. Poonen and E. Rains, "Random maximal isotropic subspaces and Selmer groups", *J. Amer. Math. Soc.* **25**:1 (2012), 245–269. MR Zbl

[Poonen and Stoll 1999] B. Poonen and M. Stoll, "The Cassels–Tate pairing on polarized abelian varieties", *Ann. of Math.* (2) **150**:3 (1999), 1109–1149. MR Zbl

[Scharlau 1985] W. Scharlau, *Quadratic and Hermitian forms*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences] **270**, Springer, 1985. MR Zbl

[Silhol 1989] R. Silhol, "Digression on real abelian varieties and classification of real abelian surfaces", Chapter 4, pp. 75–94 in *Real algebraic surfaces*, Lecture Notes in Mathematics **1392**, Springer, 1989. MR

[Smith 2016] A. Smith, "Governing fields and statistics for 4-Selmer groups and 8-class groups", 2016. arXiv

[Česnavičius 2018] K. e. Česnavičius, "The ℓ-parity conjecture over the constant quadratic extension", *Math. Proc. Cambridge Philos. Soc.* **165**:3 (2018), 385–409. MR Zbl

[Yu 2016] M. Yu, "Selmer ranks of twists of hyperelliptic curves and superelliptic curves", *J. Number Theory* **160** (2016), 148–185. MR Zbl

Communicated by Bjorn Poonen
Received 2017-12-01 Revised 2018-11-02 Accepted 2019-01-23

adam.morgan@glasgow.ac.uk *School of Mathematics and Statistics, University of Glasgow, United Kingdom*

msp

ALGEBRA AND NUMBER THEORY 13:4 (2019)
dx.doi.org/10.2140/ant.2019.13.901

Iwasawa theory for Rankin-Selberg products of p-nonordinary eigenforms

Kâzım Büyükboduk, Antonio Lei, David Loeffler and Guhan Venkat

Let f and g be two modular forms which are nonordinary at p. The theory of Beilinson–Flach elements gives rise to four rank-one nonintegral Euler systems for the Rankin–Selberg convolution $f \otimes g$, one for each choice of p-stabilisations of f and g. We prove (modulo a hypothesis on nonvanishing of p-adic L-functions) that the p-parts of these four objects arise as the images under appropriate projection maps of a single class in the wedge square of Iwasawa cohomology, confirming a conjecture of Lei–Loeffler–Zerbes.

Furthermore, we define an explicit logarithmic matrix using the theory of Wach modules, and show that this describes the growth of the Euler systems and p-adic L-functions associated to $f \otimes g$ in the cyclotomic tower. This allows us to formulate "signed" Iwasawa main conjectures for $f \otimes g$ in the spirit of Kobayashi's $\pm$-Iwasawa theory for supersingular elliptic curves; and we prove one inclusion in these conjectures under our running hypotheses.

1. Introduction

1.1. ***The setting.*** Throughout this article, we fix an odd prime p and embeddings $\iota_\infty : \overline{\mathbb{Q}} \hookrightarrow \mathbb{C}$ and $\iota_p : \overline{\mathbb{Q}} \hookrightarrow \mathbb{C}_p$. Let f and g be two normalised, new cuspidal modular eigenforms, of weights $k_f + 2$, $k_g + 2$, levels N_f, N_g, and characters ϵ_f, ϵ_g respectively. We assume that $k_f, k_g \geq 0$, that $p \nmid N_f N_g$, and that f and g are both nonordinary at p (with respect to the embeddings we fixed).

The authors' research is partially supported by: European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No. 745691 (Büyükboduk); the NSERC Discovery Grants Program 05710 (Lei, Venkat); Royal Society University Research Fellowship "L-functions and Iwasawa theory" (Loeffler); a CRM-Laval postdoctoral fellowship and a Philip Leverhulme Prize grant PLP-2014-354 funded by Leverhulme Trust (Venkat).

MSC2010: primary 11R23; secondary 11F11, 11R20.

Keywords: Iwasawa theory, elliptic modular forms, nonordinary primes.

Let $E/\mathbb{Q}_p$ be a finite extension containing the coefficients of f and g, as well as the roots of the Hecke polynomials of f and g at p. We shall write α_f, β_f, α_g and β_g for these roots; we assume throughout that $\alpha_f \neq \beta_f$ and $\alpha_g \neq \beta_g$.

Let $\mathcal{O}$ denote the ring of integers of E. For each $h \in \{f, g\}$, we fix a Galois-stable $\mathcal{O}$-lattice R_h inside Deligne's E-linear representation of $G_{\mathbb{Q}}$. The goal of this article is to study the Iwasawa theory of $T := R_f^* \otimes R_g^* = \mathrm{Hom}(R_f \otimes R_g, \mathcal{O})$ over $\mathbb{Q}(\mu_{p^\infty})$.

1.2. *Main results.*

Beilinson–Flach elements. For each of the four choices of pairs (λ, μ), where $\lambda \in \{\alpha_f, \beta_f\}$ and $\mu \in \{\alpha_g, \beta_g\}$, and each integer $m \geq 1$ coprime to p, there exists a Beilinson–Flach class

$$\mathrm{BF}_{\lambda,\mu,m} \in \mathcal{H} \otimes H^1_{\mathrm{Iw}}(\mathbb{Q}(\mu_{mp^\infty}), T),$$

as constructed in [Loeffler and Zerbes 2016b]. For the definition of $\mathcal{H}$ and H^1_{Iw}, see Section 2 below. The classes $\mathrm{BF}_{\lambda,\mu,m}$ satisfy Euler-system norm relations as m varies. However, they are not integral; that is, they do not lie $H^1_{\mathrm{Iw}}(\mathbb{Q}(\mu_{mp^\infty}), T)$. This is the chief difficulty in using these elements to study the Iwasawa theory of T.

In Section 3.4 below, we recall the relevant properties of these classes, focussing on their dependence on the choice of the p-stabilisation data (λ, μ). We also recall the explicit reciprocity law relating the Beilinson–Flach classes for $m = 1$ to p-adic Rankin–Selberg L-functions: applying the Perrin-Riou regulator map to the four classes $\mathrm{BF}_{\lambda,\mu,1}$ and projecting to suitable eigenspaces, we obtain the four unbounded p-adic L-functions associated to f and g studied in [Loeffler and Zerbes 2016b].

As a by-product of this analysis we also obtain four new p-adic L-functions, which are defined and studied in Section 3.7. We conjecture that these fall into two pairs, with each pair differing only by a sign. This gives a total of 6 p-adic L-functions for f and g, which is consistent with a conjecture of Perrin-Riou, predicting one p-adic L-function for every φ-eigenspace in the 6-dimensional space $\bigwedge^2_E \mathbb{D}_{\mathrm{cris}}(V)$.

A "rank 2" Beilinson–Flach element. In [Lei et al. 2014], it was conjectured that the four Beilinson–Flach elements associated to f and g can be seen as the images, under suitable linear functionals, of a single element in the wedge square of Iwasawa cohomology. We recall this conjecture (in a slightly strengthened form) as Conjecture 3.5.1 below. Our first main result gives a partial confirmation of this conjecture, assuming $m = 1$ and some technical hypotheses:

Theorem A. *Suppose that all four p-adic Rankin–Selberg L-functions given as in (3.6.1) are non-zero-divisors in $\mathcal{H}$, and that the following "big image" hypotheses hold*:

- *The representation V is absolutely irreducible.*
- *There exists an element $\tau \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}(\mu_{p^\infty}))$ such that the E-vector space $V/(\tau - 1)V$ is 1-dimensional.*

Then there exists a class $\mathrm{BF}_1 \in \mathrm{Frac}\,\mathcal{H} \otimes_\Lambda \bigwedge^2 H^1_{\mathrm{Iw}}(\mathbb{Q}(\mu_{p^\infty}), T)$ *whose image under the appropriate choice of Perrin-Riou functional (corresponding to the choice* $\lambda \in \{\alpha_f, \beta_f\}$ *and* $\mu \in \{\alpha_g, \beta_g\}$*) equals* $\mathrm{BF}_{\lambda,\mu,1}$.

See Theorem 3.9.1 below for a more precise formulation of Theorem A.

Decompositions using matrices of logarithms. By analogy with earlier work on Iwasawa theory of p-supersingular motives (such as [Kobayashi 2003; Lei 2011; Lei et al. 2010; 2011; Büyükboduk and Lei 2016]), one naturally expects the growth of the denominators of the Beilinson–Flach elements and p-adic L-functions for $f \otimes g$ to be governed by a suitable "matrix of logarithms", depending only on the restriction of T to the decomposition group at p.

In this paper, we construct such a logarithmic matrix for the representation T. More precisely, we show that there exists a 4×4 matrix M defined over $\mathcal{H}$ allowing us to decompose Perrin-Riou's regulator map $\mathcal{L} : H^1_{\mathrm{Iw}}(\mathbb{Q}_p, T) \to \mathcal{H} \otimes \mathbb{D}_{\mathrm{cris}}(T)$. That is, there exist four bounded Coleman maps

$$\mathrm{Col}_{\bullet,\circ} : H^1_{\mathrm{Iw}}(\mathbb{Q}_p, T) \to \mathcal{O}[\![\mathrm{Gal}(\mathbb{Q}_p(\mu_{p^\infty})/\mathbb{Q}_p)]\!], \quad \bullet, \circ \in \{\#, \flat\}$$

such that

$$\mathcal{L} = \begin{pmatrix} v_1 & v_2 & v_3 & v_4 \end{pmatrix} \cdot M \cdot \begin{pmatrix} \mathrm{Col}_{\#,\#} \\ \mathrm{Col}_{\#,\flat} \\ \mathrm{Col}_{\flat,\#} \\ \mathrm{Col}_{\flat,\flat} \end{pmatrix}$$

for some basis $\{v_1, v_2, v_3, v_4\}$ of $\mathbb{D}_{\mathrm{cris}}(T)$.

We conjecture that this logarithmic matrix can be used to decompose the unbounded Beilinson–Flach elements $\mathrm{BF}_{\lambda,\mu,m}$ into bounded classes. The precise formulation (Conjecture 5.3.1 below) is that there should exist elements $\mathrm{BF}_{\bullet,\circ,m} \in H^1_{\mathrm{Iw}}(\mathbb{Q}(\mu_{mp^\infty}), T)$ for $\bullet, \circ \in \{\#, \flat\}$ such that

$$\begin{pmatrix} \mathrm{BF}_{\alpha,\alpha,m} \\ \mathrm{BF}_{\alpha,\beta,m} \\ \mathrm{BF}_{\beta,\alpha,m} \\ \mathrm{BF}_{\beta,\beta,m} \end{pmatrix} = M \cdot \begin{pmatrix} \mathrm{BF}_{\#,\#,m} \\ \mathrm{BF}_{\#,\flat,m} \\ \mathrm{BF}_{\flat,\flat,m} \\ \mathrm{BF}_{\flat,\flat,m} \end{pmatrix}. \tag{$\dagger$}$$

We refer to these conjectural $\mathrm{BF}_{\bullet,\circ,m}$ as *doubly signed Beilinson–Flach elements*, since they depend on the two choices of symbols $(\bullet, \circ)$.

While we are currently unable to prove such a decomposition, we show a partial result in this direction in Section 5.4, where we consider the images of the unbounded Beilinson–Flach elements at locally algebraic characters in a certain range. We also show that if our Conjecture 3.5.1 on the existence of integral rank-two classes $\{\mathrm{BF}_m\}$ holds for some m, then Conjecture 5.3.1 follows as a consequence.

Bounded p-adic L-functions and Iwasawa main conjectures. On assuming that integral classes $\mathrm{BF}_{\bullet,\circ,1}$ exist satisfying ($\dagger$) for $m = 1$, we may define bounded p-adic L-functions by applying the integral Coleman maps $\mathrm{Col}_{\triangle,\square}$ to these elements. Here $(\triangle, \square) \in \{\#, \flat\}^2$ is a second pair of symbols; we refer to these p-adic L-functions as *quadruply signed*. This gives a factorisation of the unbounded p-adic

L-functions into bounded ones. We formulate an Iwasawa main conjecture (Conjecture 6.2.1 below) relating these p-adic L-functions to the characteristic ideals of the Selmer groups defined using the intersection of the kernels of $\mathrm{Col}_{\bullet,\circ}$ and $\mathrm{Col}_{\triangle,\square}$. Using the classical Euler system machine, we are able to show that one inclusion of this main conjecture holds under various technical hypotheses:

Theorem B. *Suppose that $|k_f - k_g| \geq 3$ and $p > k_f + k_g + 2$. Assume the validity of Conjecture 5.3.1, and of the hypotheses* **(A–Sym)**, **(H.nA)**, **(BI0)** *and at least one of* **(BI1)**–**(BI2)** *stated in Section 6.2 below. For any integer j with $1 + (k_f + k_g)/2 < j \leq \max(k_f, k_g)$, there exists at least one choice of symbols $\mathfrak{S} = \{(\triangle, \square), (\bullet, \circ)\}$ with $(\triangle, \square), (\bullet, \circ) \in \{\#, \flat\}^2$ such that the ω^j-isotypic component of the quadruply signed Selmer group $\mathrm{Sel}_{\mathfrak{S}}(T^\vee(1)/\mathbb{Q}(\mu_{p^\infty}))$ is $\mathcal{O}[\![\Gamma_1]\!]$-cotorsion and*

$$e_{\omega^j}\mathfrak{L}_{\mathfrak{S}} \in \mathrm{char}_{\mathcal{O}[\![\Gamma_1]\!]}\big(e_{\omega^j}\,\mathrm{Sel}_{\mathfrak{S}}(T^\vee(1)/\mathbb{Q}(\mu_{p^\infty}))^\vee\big)$$

as ideals of $\mathcal{O}[\![\Gamma_1]\!] \otimes \mathbb{Q}_p$.

See Definition 6.1.2 below where we define the quadruply signed Selmer group $\mathrm{Sel}_{\mathfrak{S}}(T^\vee(1)/\mathbb{Q}(\mu_{p^\infty}))$ and the quadruply signed p-adic L-function $\mathfrak{L}_{\mathfrak{S}}$. See also Theorem 6.2.4 for a more precise formulation of Theorem B as well as Proposition 5.3.4 where we give a sufficient condition for the validity of the condition **(A–Sym)**.

Triangulordinary Selmer groups. An alternative approach to Iwasawa theory for supersingular motives is given by Pottharst's theory of triangulordinary Selmer groups. This allows us to associate a Selmer group to each of the six φ-eigenspaces in $\mathbb{D}_{\mathrm{cris}}(V)$; these Selmer groups are finitely generated over the distribution algebra $\mathcal{H}$, rather than the Iwasawa algebra, and one expects their characteristic ideals to be generated by the associated unbounded p-adic L-functions. As a consequence of our results on the Iwasawa main conjectures for the bounded, quadruply signed p-adic L-functions, we obtain one inclusion in the Pottharst-style main conjectures, under our running hypotheses:

Theorem C (Corollary 7.4.9 below). *Suppose that all hypotheses in Theorem B hold true and choose $\mathfrak{S} = \{(\triangle, \square), (\bullet, \circ)\}$ that ensures the validity of the conclusions of Theorem B. Then for each $\lambda \in \{\alpha_f, \beta_f\}$ we have the divisibility*

$$\mathrm{char}(H^2(\mathbb{Q}, \mathbb{V}^\dagger; \mathbb{D}_\lambda)) \mid \mathrm{char}(\mathrm{coker}\,\mathrm{Col}_{\bullet,\circ})L_p(f_\lambda, g)$$

taking place in the ring $\mathcal{H}$. Here, $H^2(\mathbb{Q}, \mathbb{V}^\dagger; \mathbb{D}_\lambda)$ is the Pottharst-style Selmer group defined in Section 7.2 below and $L_p(f_\lambda, g)$ is the Rankin–Selberg p-adic L-function associated to the p-stabilisation f_λ of f.

Forthcoming work. We shall study the special case when $f = g$ with $a_p = 0$ in the subsequent article [Büyükboduk et al. 2018]. In that case, the rank-four module T decomposes into the direct sum of the symmetric square of R_f^* and a rank-one representation, and we use the results in the present paper to study the Iwasawa theory of the symmetric square of f (which is complementary to the work of Loeffler and Zerbes [2016a] in the ordinary case). In this special case, we are able to verify the nonvanishing condition that is the key hypothesis in Theorem A, allowing us to prove unconditional versions of Theorems B and C for the symmetric square.

2. Review on p-adic power series

We recall the definitions and basic properties of the rings appearing in nonordinary Iwasawa theory, following [Lei et al. 2017, §2]. We fix a finite extension $E/\mathbb{Q}_p$ with ring of integers $\mathcal{O}$, which will be the coefficient field for all the representations we shall consider.

2.1. ***Iwasawa algebras and distribution algebras.*** Let $\Gamma = \mathrm{Gal}(\mathbb{Q}(\mu_{p^\infty})/\mathbb{Q})$. This group is isomorphic to a direct product $\Delta \times \Gamma_1$, where Δ is a finite group of order $p-1$ and $\Gamma_1 = \mathrm{Gal}(\mathbb{Q}(\mu_{p^\infty})/\mathbb{Q}(\mu_p))$. We choose a topological generator γ of Γ_1, which determines an isomorphism $\Gamma_1 \cong \mathbb{Z}_p$.

We write $\Lambda = \mathcal{O}[\![\Gamma]\!]$, the Iwasawa algebra of Γ. The subalgebra $\mathcal{O}[\![\Gamma_1]\!]$ can be identified with the formal power series ring $\mathcal{O}[\![X]\!]$, via the isomorphism sending γ to $1+X$; this extends to an isomorphism

$$\Lambda = \mathcal{O}[\Delta][\![X]\!]. \tag{2.1.1}$$

We may consider Λ as a subring of the ring $\mathcal{H}$ of locally analytic E-valued distributions on Γ. The isomorphism (2.1.1) extends to an identification between $\mathcal{H}$ and the subring of power series $F \in E[\Delta][\![X]\!]$ which converge on the open unit disc $|X| < 1$.

For $n \geq 0$, we write $\omega_n(X)$ for the polynomial $(1+X)^{p^n} - 1$. We set $\Phi_0(X) = X$, and $\Phi_n(X) = \omega_n(X)/\omega_{n-1}(X)$ for $n \geq 1$. We write Tw for the ring automorphism of $\mathcal{H}$ defined by $\sigma \mapsto \chi(\sigma)\sigma$ for $\sigma \in \Gamma$. Let $u = \chi(\gamma)$ be the image of our topological generator γ under the cyclotomic character, so that Tw maps X to $u(1+X)-1$. If $m \geq 1$ is an integer, we define

$$\omega_{n,m}(X) = \prod_{i=0}^{m-1} \mathrm{Tw}^{-i}(\omega_n(X)); \qquad \Phi_{n,m}(X) = \prod_{i=0}^{m-1} \mathrm{Tw}^{-i}(\Phi_n(X))$$

Let $\log_p$ be the p-adic logarithm in $\mathcal{H}$. We define similarly

$$\log_{p,m} = \prod_{i=0}^{m-1} \mathrm{Tw}^{-i}(\log_p).$$

Finally, we define

$$\mathfrak{n}_m = \prod_{i=0}^{m-1} \mathrm{Tw}^{-i}\left(\frac{\log_p(1+X)}{X}\right).$$

2.2. ***Power series rings.*** Let $\mathbb{A}^+_{\mathbb{Q}_p} = \mathcal{O}[\![\pi]\!]$, where π is a formal variable. We equip this ring with a $\mathcal{O}$-linear *Frobenius endomorphism* φ, defined by $\pi \mapsto (1+\pi)^p - 1$, and with an $\mathcal{O}$-linear action of Γ defined by $\pi \mapsto (1+\pi)^{\chi(\sigma)} - 1$ for $\sigma \in \Gamma$, where χ denotes the p-adic cyclotomic character.

The Frobenius φ has a left inverse ψ, satisfying

$$(\varphi \circ \psi)(F)(\pi) = \frac{1}{p} \sum_{\zeta : \zeta^p = 1} F(\zeta(1+\pi) - 1).$$

The map ψ is not a morphism of rings, but it is $\mathcal{O}$-linear, and commutes with the action of Γ.

We regard $\mathbb{A}^+_{\mathbb{Q}_p}$ as a subring of the larger ring

$$\mathbb{B}^+_{\mathrm{rig},\mathbb{Q}_p} = \big\{F(\pi) \in E[\![\pi]\!] : F \text{ converges on the open unit disc}\big\}.$$

The actions of φ, ψ, and Γ extend to $\mathbb{B}^+_{\mathrm{rig},\mathbb{Q}_p}$, via the same formulae as before. We shall write $q = \varphi(\pi)/\pi \in \mathbb{A}^+_{\mathbb{Q}_p}$, and $t = \log_p(1+\pi) \in \mathbb{B}^+_{\mathrm{rig},\mathbb{Q}_p}$.

2.3. ***The Mellin transform.*** The action of Γ on $1+\pi \in (\mathbb{A}^+_{\mathbb{Q}_p})^{\psi=0}$ extends to an isomorphism of Λ-modules

$$\mathfrak{M} : \Lambda \xrightarrow{\cong} (\mathbb{A}^+_{\mathbb{Q}_p})^{\psi=0}, \quad 1 \longmapsto 1+\pi,$$

called the *Mellin transform*. This can be further extended to an isomorphism of $\mathcal{H}$-modules

$$\mathcal{H} \xrightarrow{\cong} (\mathbb{B}^+_{\mathrm{rig},\mathbb{Q}_p})^{\psi=0}$$

which we denote by the same symbol.

Theorem 2.3.1. *For all integers $m, n \geq 1$, the Mellin transform induces an isomorphism of Λ-modules*

$$\Phi_{n,m}(X)\Lambda \cong \varphi^n(q^m)(\mathbb{A}^+_{\mathbb{Q}_p})^{\psi=0}.$$

Proof. See [Lei et al. 2017, Theorem 2.1 and Equation (2.2)]. □

2.4. ***Classical and analytic Iwasawa cohomology.*** Let $\mathcal{T}$ be a finite-rank free $\mathcal{O}$-module with a continuous action of G_F, where F is a finite unramified extension of $\mathbb{Q}_p$. Then the Iwasawa cohomology groups of $\mathcal{T}$ are classically defined by

$$H^i_{\mathrm{Iw}}(F(\mu_{p^\infty}), \mathcal{T}) := \varprojlim_n H^i(F(\mu_{p^n}), \mathcal{T}).$$

Alternatively, these can be defined using a version of Shapiro's lemma: set $\mathbb{T} = \mathcal{T} \otimes \Lambda^\iota$, where Λ^ι denotes the free rank 1 Λ-module on which G_F acts via the *inverse* of the canonical character $G_F \twoheadrightarrow \Gamma \hookrightarrow \Lambda^\times$. Then one has

$$H^i_{\mathrm{Iw}}(F(\mu_{p^\infty}), \mathcal{T}) \cong H^1(F, \mathbb{T}).$$

If F is a number field, and Σ a finite set of places of F containing all $v \mid p\infty$ and all primes where $\mathcal{T}$ is ramified, then we can define similarly

$$H^i_{\mathrm{Iw},\Sigma}(F(\mu_{p^\infty}), \mathcal{T}) := \varprojlim_n H^i(F_\Sigma/F(\mu_{p^n}), \mathcal{T}) \cong H^i(F_\Sigma/F, \mathbb{T}),$$

where F_Σ is the maximal extension unramified outside Σ. In both local and global settings, the Iwasawa cohomology groups are finitely generated as Λ-modules, and zero unless $i \in \{1, 2\}$. We define Iwasawa cohomology of $\mathcal{V} = \mathcal{T}[1/p]$ by tensoring the above groups with $\mathbb{Q}_p$.

Remark 2.4.1. The group $H^1_{\mathrm{Iw},\Sigma}(F(\mu_{p^\infty}), \mathcal{T})$ is actually independent of the choice of Σ, and we will frequently drop Σ from the notation. This is not the case for $H^2_{\mathrm{Iw},\Sigma}$.

The "analytic" variants of these modules, which play a key role in Pottharst's approach [2013], to cyclotomic Iwasawa theory of nonordinary motives are obtained by systematically replacing Λ with the larger ring $\mathcal{H}$. We define $\mathbb{V}^\dagger = \mathcal{V} \otimes \mathcal{H}^\iota$; then for F a p-adic field we have

$$H^i_{\mathrm{an}}(F(\mu_{p^\infty}), \mathcal{V}) := H^i(F, \mathbb{V}^\dagger) \cong \mathcal{H} \otimes_{\Lambda[1/p]} H^i_{\mathrm{Iw}}(F(\mu_{p^\infty}), \mathcal{V}),$$

and similarly for the global setting. The importance of the analytic Iwasawa cohomology groups is that for F a p-adic field, the analytic Iwasawa cohomology of $\mathcal{V}$ is encoded in its Robba-ring (φ, Γ)-module; see Section 7.1 below.

2.5. ***The Perrin-Riou regulator map.*** Let F be an unramified extension of $\mathbb{Q}_p$, and $\mathcal{T}$ an $\mathcal{O}$-representation of G_F as before; and assume that $\mathcal{V} = \mathcal{T}[1/p]$ is crystalline, with all Hodge–Tate weights[1] ≥ 0, and with no quotient isomorphic to the trivial representation. We also fix a choice of p-power roots of unity $\zeta_{p^n} \in \overline{\mathbb{Q}}_p$, for $n \geq 1$.

Then there is a canonical homomorphism of $\mathcal{H}$-modules, the *Perrin-Riou regulator*,

$$\mathcal{L}_{F,\mathcal{V}} : H^1_{\mathrm{Iw}}(F(\mu_{p^\infty}), \mathcal{V}) \to \mathcal{H} \otimes \mathbb{D}_{\mathrm{cris}}(F, \mathcal{V})$$

which interpolates the values of the Bloch–Kato logarithm and dual-exponential maps for twists of $\mathcal{V}$ by locally algebraic characters of Γ.

It will be important to us later to consider how these maps interact with change of the field F. If K/F is an unramified extension with Galois group U, then $\mathbb{D}_{\mathrm{cris}}(K, \mathcal{V}) = K \otimes_F \mathbb{D}_{\mathrm{cris}}(F, \mathcal{V})$; so the source and target of the regulator map $\mathcal{L}_{K,\mathcal{V}}$ are naturally modules over the larger group $K(\mu_{p^\infty})/F \cong \Gamma \times U$, and it follows easily from the construction that $\mathcal{L}_{K,\mathcal{V}}$ commutes with the action of this group.

Moreover, we have an interaction with restriction and corestriction maps which can be summarised by the following diagram:

$$\begin{array}{ccc} H^1_{\mathrm{Iw}}(K(\mu_{p^\infty}), \mathcal{V}) & \xrightarrow{\mathcal{L}_{K,\mathcal{V}}} & \mathcal{H} \otimes \mathbb{D}_{\mathrm{cris}}(K, \mathcal{V}) \\ \mathrm{cores} \downarrow \quad \uparrow \mathrm{res} & & \mathrm{trace} \downarrow \quad \uparrow \subseteq \\ H^1_{\mathrm{Iw}}(F(\mu_{p^\infty}), \mathcal{V}) & \xrightarrow{\mathcal{L}_{F,\mathcal{V}}} & \mathcal{H} \otimes \mathbb{D}_{\mathrm{cris}}(F, \mathcal{V}). \end{array} \tag{2.5.1}$$

3. Euler systems of rank 2 for Rankin–Selberg products (Conjectures)

We expect that the Beilinson–Flach Euler system may be obtained by applying a suitable "rank-lowering operator" to a rank-2 Euler system. Our goal in this section is to present a precise account of this expectation and formulate a conjecture. Even though we are currently unable to verify this conjecture in general, it serves as a signpost for the signed-splitting procedure for the p-stabilised Beilinson–Flach elements that we will develop in the later sections.

[1] Our convention is that the Hodge–Tate weight of the cyclotomic character is $+1$.

3.1. ***Review of Perrin-Riou's theory.*** Let $\mathcal{T}$ be a free $\mathcal{O}$-module of finite rank with a continuous action of the absolute Galois group $G_{\mathbb{Q}}$, which is unramified outside a finite set of primes $\Sigma \ni p$. Let $\mathcal{V} = \mathcal{T} \otimes_{\mathcal{O}} E$.

Let $\mathcal{P}$ denote a set of primes disjoint from Σ, and let $\mathcal{N}(\mathcal{P})$ denote the set of square-free integers whose prime divisors are in $\mathcal{P}$. For an integer $m \in \mathcal{N}(\mathcal{P})$, we set $\Delta_m = \mathrm{Gal}(\mathbb{Q}(\mu_m)/\mathbb{Q})$ and $\Lambda_m := \mathcal{O}[\![\mathrm{Gal}(\mathbb{Q}(\mu_{mp^\infty})/\mathbb{Q})]\!] \cong \Lambda \otimes_{\mathbb{Z}_p} \mathbb{Z}_p[\Delta_m]$.

Definition 3.1.1. An Euler system of rank $r \geq 0$ is a collection of classes

$$c_m \in \bigwedge_{\Lambda_m}^{r} H^1_{\mathrm{Iw}}(\mathbb{Q}(\mu_{mp^\infty}), \mathcal{T})$$

for each $m \in \mathcal{N}(\mathcal{P})$, such that if ℓ is a prime with $\ell, m\ell \in \mathcal{N}(\mathcal{P})$, then

$$\mathrm{cor}_{\mathbb{Q}(\mu_{m\ell p^\infty})/\mathbb{Q}(\mu_{mp^\infty})}(c_{m\ell}) = \begin{cases} P_\ell(\sigma_\ell^{-1})c_m & \text{if } \ell \nmid m, \\ c_m & \text{if } \ell \mid m. \end{cases} \tag{3.1.1}$$

Here $P_\ell(X) := \det_E(1 - \mathrm{Frob}_\ell^{-1} X \mid V^*(1)) \in \mathcal{O}[X]$, and σ_ℓ denotes the image in $\mathrm{Gal}(\mathbb{Q}(\mu_{mp^\infty})/\mathbb{Q})$ of Frob_ℓ, the arithmetic Frobenius at ℓ.

Remark 3.1.2. Perrin-Riou in fact requires $r \geq 1$, but we feel that the case $r = 0$ should not be neglected. For $r = 0$ we have $\bigwedge^r_{\Lambda_m} H^1_{\mathrm{Iw}}(\mathbb{Q}(\mu_{mp^\infty}), \mathcal{T}) = \Lambda_m$, and a rank 0 Euler system is therefore a collection of elements $c_m \in \Lambda_m$ for $m \in \mathcal{N}(\mathcal{P})$, satisfying the compatibilities (3.1.1) under the projection maps $\Lambda_{m\ell} \to \Lambda_m$. Such collections of elements arise naturally in the theory of p-adic L-functions: for instance, both the Stickelberger elements for an odd Dirichlet character, and the Mazur–Tate elements for a p-ordinary modular form, can be viewed as rank 0 Euler systems in this sense.

Given an Euler system of rank $r > 1$, one can construct a multitude of Euler systems of rank 1 with the aid of auxiliary choices of functionals on the Iwasawa cohomology, following a recipe originally set out by Rubin [1996] and later formalised by Perrin-Riou [1998].

Definition 3.1.3. A *Perrin-Riou functional* of rank $s \geq 1$ is a collection of linear functionals

$$\{\Phi_m : m \in \mathcal{N}(\mathcal{P})\},$$

where

$$\Phi_m \in \bigwedge_{\Lambda_m}^{s} \mathrm{Hom}_{\Lambda_m}\big(H^1_{\mathrm{Iw}}(\mathbb{Q}(\mu_{mp^\infty}), \mathcal{T}), \Lambda_m\big),$$

such that if ℓ is a prime with $\ell, m\ell \in \mathcal{N}(\mathcal{P})$, we have

$$\Phi_{m\ell} \circ \mathrm{res}_{\mathbb{Q}(\mu_{m\ell p^\infty})/\mathbb{Q}(\mu_{mp^\infty})} = \iota_{m\ell/m} \circ \Phi_m, \tag{3.1.2}$$

where $\iota_{ml/m}$ denotes the isomorphism $\Lambda_m \cong (\Lambda_{m\ell})^{\Delta_\ell}$ sending 1 to $\sum_{\sigma \in \Delta_\ell} [\sigma]$.

As in [Rubin 1996, Corollary 1.3], one may interpret a rank $r-1$ Perrin-Riou functional $\Phi = \{\Phi_m\}$ as a collection of maps

$$\Phi_m : \bigwedge_{\Lambda_m}^{r} H^1_{\mathrm{Iw}}(\mathbb{Q}(\mu_{mp^\infty}), \mathcal{T}) \to H^1_{\mathrm{Iw}}(\mathbb{Q}(\mu_{mp^\infty}), \mathcal{T}).$$

Proposition 3.1.4 (Perrin-Riou). *If $\{c_m\}_{m\in N(\mathcal{P})}$ is an Euler system of rank r and $\{\Phi_m\}$ is a Perrin-Riou functional of rank $r-1$, then*

$$\Phi_m(c_m) \in H^1_{\mathrm{Iw}}(\mathbb{Q}(\mu_{mp^\infty}), \mathcal{T})$$

is a rank one Euler system.

Proof. See [Perrin-Riou 1998, Lemma 1.2.3; Rubin 1996, §6]. □

Remark 3.1.5. More generally, one may interpret a rank s Perrin-Riou functional as a "rank-lowering operator" sending rank r Euler systems to rank $r-s$ Euler systems. This includes the case $r = s$, where we understand rank 0 Euler systems as in Remark 3.1.2 above.

3.2. ***Analytic Euler systems.*** For the applications below, we will need to consider compatible families of classes not lying in Iwasawa cohomology, but in the larger "analytic" cohomology modules of Pottharst. For simplicity, we shall only describe this construction in rank 1.

We recall that $\mathcal{H}$ can be written as an inverse limit $\varprojlim_n \mathcal{H}[n]$, where $\mathcal{H}[n]$ are reduced affinoid algebras, and for each n we have

$$\mathcal{H}[n] \otimes_{\mathcal{H}} H^1_{\mathrm{an}}(\mathbb{Q}(\mu_{mp^\infty}), \mathcal{V}) = H^1(\mathbb{Q}(\mu_m), \mathcal{H}_n \otimes_E \mathcal{V})$$

by [Pottharst 2013, Theorem 1.7]. Each $\mathcal{H}[n]$ has a canonical supremum norm; if $\mathcal{H}[n]^\circ$ denotes the unit ball for this norm, then there is a seminorm $\|\cdot\|_n$ on $H^1(\mathbb{Q}(\mu_m), \mathcal{H}[n] \otimes_E \mathcal{V})$ for which the unit ball is the image of $H^1(\mathbb{Q}(\mu_m), \mathcal{H}[n]^\circ \otimes_{\mathcal{O}} \mathcal{T})$. (In fact this is a norm, by [Loeffler and Zerbes 2016b, Proposition 2.1.2(1)], but we do not need this.)

Definition 3.2.1. An *analytic Euler system* (of rank 1) for $\mathcal{V}$ is a collection of classes $c_m \in H^1_{\mathrm{an}}(\mathbb{Q}(\mu_{mp^\infty}), \mathcal{V})$, for each $m \in \mathcal{N}(\mathcal{P})$, satisfying the following two conditions:

(1) if ℓ is prime and $m, m\ell \in \mathcal{N}(\mathcal{P})$, then the norm-compatibility condition (3.1.1) holds;

(2) for each n, there is a constant C_n (independent of m) such that $\|c_m\|_n \le C_n$ for all $m \in \mathcal{N}(\mathcal{P})$.

Remark 3.2.2. Condition (1), asserting that there is no "growth in the tame direction", is technical to state but absolutely vital in order to obtain an interesting theory; it is trivial that any class in $H^1_{\mathrm{an}}(\mathbb{Q}(\mu_{p^\infty}), \mathcal{V})$ can be extended to a compatible family of classes satisfying (2) alone.

3.3. ***Unbounded Perrin-Riou functionals.*** In this section, building on [Otsuki 2009; Lei et al. 2014], we will construct *canonical* Perrin-Riou functionals using another construction of Perrin-Riou, namely the p-adic regulator map. The price we pay for this canonicity is that our functionals are no longer bounded in general.

We now assume $\mathcal{V}$ is crystalline, with all Hodge–Tate weights ≥ 0, and that $\mathcal{V}|_{G_{\mathbb{Q}_p}}$ has no quotient isomorphic to the trivial representation. We have already chosen a compatible family of p-power roots of unity ζ_{p^r}. For $\mathcal{P}$ as above, let us also choose a primitive n-th root of unity ζ_ℓ for each $\ell \in \mathcal{P}$. One checks easily that if $m \in \mathcal{N}(\mathcal{P})$, then $\xi_m = \prod_{\ell|m}(-\zeta_\ell)$ is a basis vector of the ring of integers $\mathbb{Z}[\mu_m]$ as a free rank 1 module over the group ring $\mathbb{Z}[\Delta_m]$; and we have the trace-compatibility

$$\mathrm{trace}_{m\ell/m}(\xi_{m\ell}) = \xi_m.$$

Definition 3.3.1. For $m \in \mathcal{N}(\mathcal{P})$, let $\nu_m : \mathbb{Z}[\zeta_m] \to \mathbb{Z}[\Delta_m]$ denote the unique $\mathbb{Z}[\Delta_m]$-linear map sending ξ_m to 1.

Let us now set $\mathcal{H}_m = \mathbb{Z}_p[\Delta_m] \otimes \mathcal{H}$, which we regard as an "analytification" of Λ_m. For $\mathcal{V}$ as above, the sum of the Perrin-Riou regulators at the primes of $\mathbb{Q}(\mu_m)$ above p gives a map

$$H^1_{\mathrm{Iw}}(\mathbb{Q}(\mu_{mp^\infty}) \otimes \mathbb{Q}_p, \mathcal{V}) \to \mathbb{Q}(\mu_m) \otimes_{\mathbb{Q}} \mathcal{H} \otimes_E \mathbb{D}_{\mathrm{cris}}(\mathbb{Q}_p, \mathcal{V}),$$

and composing this with ν_m we obtain a morphism of $\mathcal{H}_m$-modules

$$\mathcal{L}_{m,\mathcal{V}} : H^1_{\mathrm{Iw}}(\mathbb{Q}(\mu_{mp^\infty}) \otimes \mathbb{Q}_p, \mathcal{V}) \to \mathcal{H}_m \otimes \mathbb{D}_{\mathrm{cris}}(\mathbb{Q}_p, \mathcal{V}). \tag{3.3.1}$$

(We shall abbreviate $\mathcal{L}_{1,\mathcal{V}}$ by $\mathcal{L}_V$.)

Definition 3.3.2. If $t \in \mathbb{D}_{\mathrm{cris}}(\mathbb{Q}_p, \mathcal{V}^*(1))$, and $m \in \mathcal{N}(\mathcal{P})$, then we define a map

$$\Phi^{(t)}_m : H^1_{\mathrm{Iw}}(\mathbb{Q}(\mu_{mp^\infty}), \mathcal{V}) \to \mathcal{H}_m, \quad \Phi^{(t)}_m(z) := \big\langle \mathcal{L}_{m,\mathcal{V}}(\mathrm{loc}_p\, z), t \big\rangle.$$

One sees easily that, for any fixed t, the collection $\Phi^{(t)} = \{\Phi^{(t)}_m : m \in \mathcal{N}(\mathcal{P})\}$ satisfies the compatibility condition (3.1.2), and thus may be regarded as a (rank 1) *unbounded Perrin-Riou functional*. Pairing with $\Phi^{(t)}$ therefore defines a homomorphism from Euler systems of rank 2 to (possibly unbounded) analytic Euler systems of rank 1.

Remark 3.3.3. Via exactly the same construction, for any $s \geq 1$ we may use elements of the wedge power $\bigwedge^s_E \mathbb{D}_{\mathrm{cris}}(\mathbb{Q}_p, \mathcal{V}^*(1))$ to define unbounded Perrin-Riou functionals of rank s.

3.4. ***The Beilinson–Flach Euler systems.*** We now focus on the particular case which interests us: the Rankin–Selberg convolution of two modular forms. As in the introduction, f and g denote two normalised, new cuspidal modular eigenforms, of weights $k_f + 2, k_g + 2$, levels N_f, N_g, and characters ϵ_f, ϵ_g respectively. We assume that $p \nmid N_f N_g$. We let T denote the rank 4 representation $R^*_f \otimes R^*_g$ over $\mathcal{O}$ and write $V = T \otimes_{\mathbb{Z}_p} \mathbb{Q}_p$.

We take Σ to be the set of primes dividing pN_fN_g, and for $\ell \notin \Sigma$, we write

$$P_\ell(X) = \det\big(1 - \mathrm{Frob}_\ell^{-1}\, X \;\big|\; T^*(1)\big) = 1 - \frac{a_\ell(f)a_\ell(g)}{\ell} X + \cdots \in \mathcal{O}[X].$$

We now review the construction of Beilinson–Flach elements from [Loeffler and Zerbes 2016b]. For $\lambda \in \{\alpha, \beta\}$ and $h \in \{f, g\}$, we write h^λ for the p-stabilisation of h at λ_h. We shall identify R^*_h

with $R^*_{h^\lambda}$ following [op. cit., §3.5]. More specifically, let pr_1 and pr_2 be the two degeneracy maps on the modular curves $Y_1(pN_h) \to Y_1(N_h)$ as defined in [Kings et al. 2017, Definition 2.4.1] and write $\mathrm{Pr}^\lambda = \mathrm{pr}_1 - (\lambda'/p^{k_h+1})\mathrm{pr}_2$, where λ' denotes the unique element of $\{\alpha, \beta\} \setminus \{\lambda\}$. Realising $R^*_{h^\lambda}$ and R^*_h as quotients of the étale cohomology of $Y_1(pN_h)$ and $Y_1(N_h)$ respectively, Pr^λ_* gives an isomorphism between these two Galois representations.

Definition 3.4.1. For $\lambda, \mu \in \{\alpha, \beta\}$, $c > 1$ coprime to $6pN_fN_g$, $m \geq 1$ coprime to pc, and $a \in (\mathbb{Z}/mp^\infty\mathbb{Z})^\times$, let

$$ {}_c\mathcal{BF}^{\lambda,\mu}_{m,a} \in D_{\mathrm{ord}_p(\lambda_f\mu_g)}(\Gamma) \otimes_\Lambda H^1_{\mathrm{Iw}}(\mathbb{Q}(\mu_{mp^\infty}), T) $$

be the Beilinson–Flach element as constructed in [Loeffler and Zerbes 2016b, Theorem 5.4.2].

Here $D_{\mathrm{ord}_p(\lambda_f\mu_g)}(\Gamma, E)$ denotes the Λ-submodule of $\mathcal{H}$ consisting of tempered distributions of order $\mathrm{ord}_p(\lambda_f\mu_g)$. We shall take $a = 1$ throughout, and restrict to integers $m \in \mathcal{N}(\mathcal{P})$, where $\mathcal{P}$ is the set of primes not dividing pcN_fN_g.

Remark 3.4.2. If $\epsilon_f\epsilon_g$ is nontrivial, then we may remove the dependence on the auxiliary integer c, but this will not greatly concern us here: we shall simply fix a value of c and drop it from the notation.

These elements satisfy a norm-compatibility relation which is close, but not identical, to Equation (3.1.1). As explained in [Lei et al. 2014, Lemma 7.3.2], we can modify these elements to "correct" the norm relation: there exists a collection of elements $\mathrm{BF}_{\lambda,\mu,m}$ for $m \in \mathcal{N}(\mathcal{P})$ such that

- the $\mathrm{BF}_{\lambda,\mu,m}$ for m varying satisfy Equation (3.1.1) exactly,
- $\mathrm{BF}_{\lambda,\mu,1} = {}_c\mathcal{BF}^{\lambda,\mu}_{1,1}$,
- each $\mathrm{BF}_{\lambda,\mu,m}$ is an $\mathcal{O}[\Delta_m]$-linear combination of the elements ${}_c\mathcal{BF}^{\lambda,\mu}_{m',1}$ for $m' \mid m$.

As in [Loeffler and Zerbes 2016b, Theorem 8.1.4(ii)], if $H^0(\mathbb{Q}(\mu_{p^\infty}), V) = 0$, the collection of elements $\mathrm{BF}_{\lambda,\mu,m}$ for varying $m \in \mathcal{N}(\mathcal{P})$ form an analytic Euler system in the sense of Definition 3.2.1. We thus obtain four rank 1 analytic Euler systems for T, one for each of the possible choices of p-stabilisations λ, μ.

One of the key themes in the present paper will be to understand the relations among these Euler systems, for a fixed f and g and different choices of p-stabilisations. Our first result in this direction is the following straightforward compatibility. For χ any continuous character of Γ, and $z \in \mathcal{H} \otimes H^1_{\mathrm{Iw}}(\mathbb{Q}(\mu_{mp^\infty}), V)$, let us write $z(\chi)$ for the image of z in $H^1(\mathbb{Q}(\mu_m), V(\chi^{-1}))$.

Lemma 3.4.3. *Let $m \in \mathcal{N}(\mathcal{P})$, and let χ be a character of Γ of the form $z \mapsto z^j\theta(z)$, where $j \in \mathbb{Z}$ and θ is a finite-order character of conductor p^r.*

(i) *If $0 \leq j \leq \min(k_f, k_g)$, then*

$$ (\lambda_f\mu_g)^r \cdot \mathrm{BF}_{\lambda,\mu,m}(\chi) $$

is independent of the choice of λ and μ.

(ii) *If $k_g < j \leq k_f$ then this class is independent of λ (but may depend on μ); and similarly if $k_f < j \leq k_g$ it is independent of μ.*

If χ is the character $z \mapsto z^j$, the same conclusions hold for the class

$$\left(1 - \frac{\lambda\mu}{p^{1+j}\sigma_p}\right)\left(1 - \frac{p^j\sigma_p}{\lambda\mu}\right)^{-1} \mathrm{BF}_{\lambda,\mu,m}(\chi).$$

Proof. Part (i) follows from the same proof as [Büyükboduk and Lei 2016, Proposition 3.3 and Corollary 3.4] since the infinite part of χ (the character $z \mapsto z^j$) has the effect of sending the Beilinson–Flach element for the representation T to that for the Tate twist $T(-j)$. For part (ii), we assume $k_g < j \leq k_f$ without loss of generality, and deform g_μ in a Coleman family $\mathcal{G}$ (while keeping f and θ fixed). By Theorem A of [Loeffler and Zerbes 2016b], we obtain two families of cohomology classes $\mathrm{BF}_{\alpha,\mathcal{G},m}(\chi)$ and $\mathrm{BF}_{\beta,\mathcal{G},m}(\chi)$, and by part (i) the specialisations of these at integer points $r' \geq j$ are equal. Hence the two families of classes are equal identically, and we obtain (ii) by specialising back to g_μ. □

3.5. ***A conjectural rank 2 Euler system.*** Recall that T is a free $\mathcal{O}$-module of rank 4 and observe that both -1 and $+1$-eigenspaces for the action of complex conjugation on T have rank 2. In this situation, we expect to have an Euler system of rank $r = 2$.

We are now ready to state our conjecture on the relation of p-stabilised Beilinson–Flach classes and rank-2 Euler systems. Let $\mathcal{L}_{m,V} : H^1_{\mathrm{Iw}}(\mathbb{Q}(\mu_{mp^\infty}), V) \to \mathcal{H}_m \otimes \mathbb{D}_{\mathrm{cris}}(V)$ be the equivariant Perrin-Riou regulator as in Equation (3.3.1), and let $\{v_{\lambda\mu}\}_{\lambda,\mu\in\{\alpha,\beta\}}$ be an eigenvector basis of $\mathbb{D}_{\mathrm{cris}}(V)$ in which the matrix of φ is given by

$$D := \begin{pmatrix} \frac{1}{\alpha_f\alpha_g} & & & \\ & \frac{1}{\alpha_f\beta_g} & & \\ & & \frac{1}{\beta_f\alpha_g} & \\ & & & \frac{1}{\beta_f\beta_g} \end{pmatrix}.$$

Further, let $\{v^*_{\lambda\mu}\}$ be the dual basis to $\{v_{\lambda\mu}\}$.

These vectors are *a priori* only determined up to scaling; we may normalise them canonically as follows. The 1-dimensional space $\mathbb{D}_{\mathrm{cris}}(V_f)/\mathrm{Fil}^1$ has a canonical basis vector η'_f, as defined in [Kings et al. 2015, §6.1]. Since we are assuming f to be nonordinary, the two eigenspaces are both complementary to Fil^1, and we can thus define $v_{f,\alpha}$ and $v_{f,\beta}$ to be the unique vectors in the φ-eigenspaces satisfying

$$v_{f,\alpha} = v_{f,\beta} = \eta'_f \mod \mathrm{Fil}^1.$$

Defining $v_{g,\mu}$ analogously, we can choose our eigenvector basis of $\mathbb{D}_{\mathrm{cris}}(V^*) = \mathbb{D}_{\mathrm{cris}}(V_f) \otimes \mathbb{D}_{\mathrm{cris}}(V_g)$ by setting $v^*_{\lambda\mu} = v_{f,\lambda} \otimes v_{g,\mu}$.

Conjecture 3.5.1. *There exists a collection of classes $\mathrm{BF}_m \in \bigwedge^2 H^1_{\mathrm{Iw}}(\mathbb{Q}(\mu_m), T)$, for all $m \in \mathcal{N}(\mathcal{P})$, which form a rank 2 Euler system, and are such that for all $\lambda, \mu \in \{\alpha, \beta\}$, we have*

$$\left\langle \mathcal{L}_{m,V}(\mathrm{BF}_m), v^*_{\lambda,\mu} \right\rangle = \mathrm{BF}_{\lambda,\mu,m}.$$

Equivalently, the four rank 1 analytic Euler systems $(\mathrm{BF}_{\lambda,\mu,m})_{m\in\mathcal{N}(\mathcal{P})}$, for different choices of λ and μ, are all obtained from the single rank 2 Euler system (BF_m) via the Perrin-Riou functionals associated to the four eigenvectors $v^*_{\lambda,\mu}$.

Remark 3.5.2. This conjecture is an extension to higher-weight modular forms of the conjectures formulated in [Lei et al. 2014, §8] for pairs of weight 2 modular forms. At the time, this conjecture was somewhat tentative since the methods of [op. cit.] only suffice to construct the classes $\mathrm{BF}_{\lambda,\mu,m}$ when $v_p(\lambda\mu) < 1$, which is satisfied for at most two of the four possible choices, and sometimes for none at all. However, this restriction has since been removed in [Loeffler and Zerbes 2016b] via the use of Coleman families.

3.6. ***p-adic L-functions and explicit reciprocity laws.*** For $\lambda \in \{\alpha_f, \beta_f\}$ and $\mu \in \{\alpha_g, \beta_g\}$, there exist Coleman families $\mathcal{F}$ and $\mathcal{G}$ passing through the p-stabilisations f_λ and g_μ; these are families of overconvergent eigenforms over some affinoid discs V_1 and V_2 in the weight space $\mathcal{W}$. We suppose (temporarily) that our coefficient field E contains a primitive N-th root of unity, where $N = \mathrm{LCM}(N_f, N_g)$.

Theorem 3.6.1 [Loeffler and Zerbes 2016b]. *There exists a 3-variable p-adic L-function $L_p^{\mathrm{geom}}(\mathcal{F},\mathcal{G}) \in \mathcal{O}(V_1\times V_2\times\mathcal{W})$ with the following interpolation property. Let (r, r', j) be an integer point in $V_1\times V_2\times\mathcal{W}$ such that $r \geq 0$, $r' \geq -1$ and $(r+r'+1)/2 \leq j \leq r$. Suppose that the specialisations $\mathcal{F}_r$ and $\mathcal{G}_{r'}$ are p-stabilisations of classical newforms f_r and $g_{r'}$ of prime-to-p level. Then,*

$$L_p^{\mathrm{geom}}(\mathcal{F},\mathcal{G})(r,r',j) = \frac{\mathcal{E}(f_r, g_{r'}, 1+j)}{\mathcal{E}(f_r)\mathcal{E}^*(f_r)} \frac{j!(j-r'-1)!\big(c^2 - c^{2j-r-r'}\epsilon_{\mathcal{F}}(c)^{-1}\epsilon_{\mathcal{G}}(c)^{-1}\big)}{\pi^{2j-r'+1}(-1)^{r-r'}2^{2j+2+r-r'}} \frac{L(f_r, g_{r'}, 1+j)}{\langle f_r, f_r\rangle_{N_f}},$$

where

$$\mathcal{E}(f_r) = \left(1 - \frac{\lambda'_r}{p\lambda_r}\right), \quad \mathcal{E}^*(f_r) = \left(1 - \frac{\lambda'_r}{\lambda_r}\right),$$
$$\mathcal{E}(f_r, g'_r, 1+j) = \left(1 - \frac{p^j}{\lambda_r\mu_r}\right)\left(1 - \frac{p^j}{\lambda_r\mu'_r}\right)\left(1 - \frac{\lambda'_r\mu_r}{p^{1+j}}\right)\left(1 - \frac{\lambda'_r\mu'_r}{p^{1+j}}\right).$$

Here, λ_r, $\mu_{r'}$ are the respective specialisations of the U_p-eigenvalues on $\mathcal{F}$ and $\mathcal{G}$ at r and r', whereas λ'_r and μ'_r are defined by the requirement that $\{\lambda_r, \lambda'_r\} = \{\alpha_{f_r}, \beta_{f_r}\}$ and $\{\mu, \mu'\} = \{\alpha_{g_{r'}}, \beta_{g_{r'}}\}$.

Remark 3.6.2. The construction of this function in [Loeffler and Zerbes 2016b] relies on deforming Beilinson–Flach elements in families. An alternative, more direct construction (not using Euler systems) has subsequently been given by Urban [2017].

Proposition 3.6.3. *Let $L_p(\mathcal{F}, g_\mu)$ denote the function on $V_1\times\mathcal{W}$ obtained by specialising $L_p^{\mathrm{geom}}(\mathcal{F},\mathcal{G})$ at the point of V_2 corresponding to g_μ. Then the functions $L_p(\mathcal{F}, g_\alpha)$ and $L_p(\mathcal{F}, g_\beta)$ coincide.*

Proof. From the preceding theorem, one sees that these two functions agree at all points (r, j) with r, j integers satisfying the inequalities $r \geq 0$ and $(r+k_g+1)/2 \leq j \leq r$. These points are clearly Zariski-dense in $V_1\times\mathcal{W}$. □

Definition 3.6.4. For $\lambda \in \{\alpha_f, \beta_f\}$, define $L_p(f_\lambda, g) \in \mathcal{O}(\mathcal{W})$ to be the specialisation of

$$\left[w(f)G(\epsilon_f^{-1})G(\epsilon_g^{-1})\mathcal{E}(f)\mathcal{E}^*(f)\right] \cdot L_p(\mathcal{F}, g)$$

at the point f_λ of V_1, where $L_p(\mathcal{F}, g)$ is the common value $L_p(\mathcal{F}, g_\alpha) = L_p(\mathcal{F}, g_\beta)$, $G(\cdot)$ are the Gauss sums, and $w(f)$ is the Atkin–Lehner pseudo-eigenvalue of f.

One can check that $L_p(f_\lambda, g)$ is defined over any p-adic field containing the coefficients of f_α and g (not necessarily containing an N-th root of unity); this is the reason for renormalising by the Gauss sums. Since there is a canonical isomorphism $\mathcal{O}(\mathcal{W}) \cong \mathcal{H}$, we shall regard $L_p(f_\lambda, g)$ as an element of $\mathcal{H}$. We therefore have four p-adic L-functions attached to the pair $\{f, g\}$, namely

$$\left\{L_p(f_\alpha, g),\ L_p(f_\alpha, g),\ L_p(g_\alpha, f),\ L_p(g_\beta, f)\right\}. \tag{3.6.1}$$

Theorem 3.6.5 (explicit reciprocity law). *For each pair (λ, μ) we have*

$$\begin{aligned}
\left\langle \mathcal{L}_V(\mathrm{BF}_{\lambda,\mu,1}), v^*_{\lambda,\mu}\right\rangle &= 0,\\
\left\langle \mathcal{L}_V(\mathrm{BF}_{\lambda,\mu,1}), v^*_{\lambda,\mu'}\right\rangle &= \frac{A_g \log_{p,1+k_g}}{(\mu'-\mu)} \cdot L_p(f_\lambda, g),\\
\left\langle \mathcal{L}_V(\mathrm{BF}_{\lambda,\mu,1}), v^*_{\lambda',\mu}\right\rangle &= \frac{A_f \log_{p,1+k_f}}{(\lambda'-\lambda)} \cdot L_p(g_\mu, f)
\end{aligned}$$

where A_f and A_g are nonzero constants independent of λ and μ. In particular we have the antisymmetry relations

$$\left\langle \mathcal{L}_V(\mathrm{BF}_{\lambda,\mu,1}), v^*_{\lambda,\mu'}\right\rangle = -\left\langle \mathcal{L}_V(\mathrm{BF}_{\lambda,\mu',1}), v^*_{\lambda,\mu}\right\rangle, \quad \left\langle \mathcal{L}_V(\mathrm{BF}_{\lambda,\mu,1}), v^*_{\lambda',\mu}\right\rangle = -\left\langle \mathcal{L}_V(\mathrm{BF}_{\lambda',\mu,1}), v^*_{\lambda,\mu}\right\rangle.$$

Proof. The vanishing of $\left\langle \mathcal{L}_V(\mathrm{BF}_{\lambda,\mu}), v^*_{\lambda,\mu}\right\rangle$ is a consequence of Theorem 7.1.2 of [Loeffler and Zerbes 2016b]. The other two formulae follow directly from the definition of the geometric p-adic L-function [op. cit., Definition 9.1.1] after a somewhat tedious comparison of conventions. The factor $\log_{p,1+k_g}$ arises from the normalisation of the Perrin-Riou regulator for a certain subquotient of the (φ, Γ)-module of V [op. cit., Theorem 7.1.4]. The quantity $A_g/(\mu'-\mu)$ and its cousin arise from comparing the families of eigenvectors constructed there with our present conventions; some handle-turning shows that the specialisation of the family $\eta_{\mathcal{F}}$ in their notation corresponds to

$$\frac{1}{w(f)G(\epsilon_f^{-1})\mathcal{E}(f_\lambda)\mathcal{E}^*(f_\lambda)} v_{f,\alpha},$$

while the family $\omega_{\mathcal{G}}$ specialises to

$$\frac{(-1)^{k_g}\langle \varphi(\omega'_g), \omega'_{g^*}\rangle}{(\mu'-\mu)N_{\epsilon_g}G(\epsilon_g^{-1})} v_{g,\beta},$$

where ω'_g is the basis vector of $\mathrm{Fil}^1\, \mathbb{D}_{\mathrm{cris}}(R_g)$ defined in [Kings et al. 2015, §6.1], and ω'_{g^*} its analogue for the conjugate form g^*. □

3.7. *Some "extra" p-adic L-functions.*

Definition 3.7.1. For $\lambda \in \{\alpha_f, \beta_f\}$ and $\mu \in \{\alpha_g, \beta_g\}$, we set

$$L_p^?(f_\lambda, g_\mu) := \frac{1}{\log_{p,\nu+1}} \langle \mathcal{L}_V(\mathrm{BF}_{\lambda,\mu,1}), v^*_{\lambda',\mu'} \rangle,$$

where $\nu = \min(k_f, k_g)$.

These elements lie in $\mathcal{H}$, because $\mathrm{BF}_{\lambda,\mu,1}(\chi)$ is in H^1_{f} for every locally algebraic χ of weight in the range $[0, \ldots, \nu]$, so $\mathcal{L}_V(\mathrm{BF}_{\lambda,\mu,1})$ vanishes at these characters and thus is divisible by $\log_{p,\nu+1}$.

Proposition 3.7.2. *Suppose $k_f > k_g$, and let χ be a character of Γ of the form $z \mapsto z^j \theta(z)$, where $k_g + 1 \leq j \leq k_f$ and θ is a Dirichlet character of conductor p^n. Then we have*

$$L_p^?(f_\lambda, g_\mu)(\chi) = R \cdot \frac{A_g}{\mu' - \mu} \cdot L_p(f_{\lambda'}, g)(\chi),$$

where A_g is as in the statement of Theorem 3.6.5; $R = (\lambda'/\lambda)^n$ if $n \geq 1$, and if $n = 0$ then

$$R = \left(1 - \frac{\lambda'\mu}{p^{1+j}\sigma_p}\right)\left(1 - \frac{p^j\sigma_p}{\lambda'\mu}\right)^{-1} \Big/ \left(\left(1 - \frac{\lambda\mu}{p^{1+j}\sigma_p}\right)\left(1 - \frac{p^j\sigma_p}{\lambda\mu}\right)^{-1}\right).$$

Proof. Applying $\langle \mathcal{L}_V(\,\cdot\,), v^*_{\lambda',\mu'} \rangle(\chi)$ to both $\mathrm{BF}_{\lambda,\mu,1}$ and $\mathrm{BF}_{\lambda',\mu,1}$, we deduce that

$$\begin{aligned} L_p^?(f_\lambda, g_\mu)(\chi) &= \frac{\langle \mathcal{L}_V(\mathrm{BF}_{\lambda,\mu,1}), v^*_{\lambda',\mu'} \rangle(\chi)}{\log_{p,\nu+1}} \\ &= R \cdot \frac{\langle \mathcal{L}_V(\mathrm{BF}_{\lambda',\mu,1}), v^*_{\lambda',\mu'} \rangle(\chi)}{\log_{p,\nu+1}} \\ &= R \cdot \frac{A_g}{\mu' - \mu} L_p(f_{\lambda'}, g)(\chi), \end{aligned}$$

where the second equality follows from Lemma 3.4.3, the final equality from Theorem 3.6.5 and the first from definitions. □

Note that for $n = 0$ the right-hand side is an explicit multiple of a complex L-value, by the explicit reciprocity law (and we expect this also to hold for $n \geq 1$). This construction gives rise to four extra elements associated to f and g, in addition to the more familiar four given by (3.6.1). It seems natural to conjecture that

$$L_p^?(f_\lambda, g_\mu) = -L_p^?(f_{\lambda'}, g_{\mu'}), \tag{3.7.1}$$

so that these four extra elements fall into two pairs differing by signs; this would, for instance, follow easily from Conjecture 3.5.1. However, we do not know how to prove this symmetry property unconditionally, since these elements do not seem to deform in Coleman families, and their growth (which is always $O(\log_{p,1+\max(k_f,k_g)})$) is just too rapid for the interpolating property to imply (3.7.1).

Remark 3.7.3. In the analogous case when f is supersingular but g is an ordinary CM form, these extra L-functions correspond to the extra two p-adic L-functions constructed in [Loeffler 2014] using modular symbols for Bianchi groups.

3.8. *Nontriviality of Beilinson–Flach elements.*

Corollary 3.8.1. *If $|k_f - k_g| \geq 3$ then for each choice of λ and μ, the class $\mathrm{res}_p(\mathrm{BF}_{\lambda,\mu,1})$ is nontrivial.*

Proof. By symmetry, we may suppose that $k_f - k_g \geq 3$. Notice that the Euler product for the Rankin–Selberg L-series $L(f, g, s)$ converges absolutely at $s = k_f + 1$ (since $k_f + 1 > (k_f + k_g)/2 + 2$). In particular, $L(f, g, k_f + 1)$ is nonzero. For either value $\lambda \in \{\alpha_f, \beta_f\}$, the factors $(c^2 - \cdots)$ and $\mathcal{E}(f, g, 1 + k_f)$ appearing in Theorem 3.6.1 are easily seen to be nonzero as well, using the fact that α_f, β_f are Weil numbers of weight $(k_f + 1)/2 > (k_g + 1)/2 + 1$, whereas α_g, β_g are Weil numbers of weight $(k_g + 1)/2$. Hence $L_p(f_\lambda, g)(k_f)$ is nonzero for both values of λ. By the explicit reciprocity law, this forces all four elements $\mathrm{res}_p(\mathrm{BF}_{\lambda,\mu,1})$ to be nonzero. □

Definition 3.8.2. For a character η of $\Delta = \Gamma_{\mathrm{tor}}$, we let $e_\eta \in \Lambda$ denote the corresponding idempotent.

Remark 3.8.3. One may show that the projection $\mathrm{res}_p(e_{\omega^j}\mathrm{BF}_{\lambda,\mu,1})$ is nontrivial for $1 + (k_f + k_g)/2 < j \leq \max(k_f, k_g)$, by arguing as in the proof of Corollary 3.8.1.

Remark 3.8.4. Note that the interpolation formula for $L_p^{\mathrm{geom}}(\mathcal{F}, \mathcal{G})$ we have recorded in Theorem 3.6.1 does not say anything about its value at $(r, r', j + \chi)$ where χ is a nontrivial finite order character of p-power conductor. This is the reason why we assume $|k_f - k_g| \geq 3$ in Corollary 3.8.1; with a stronger interpolation formula we could reduce this to $|k_f - k_g| \geq 2$, and even $|k_f - k_g| \geq 1$ conditionally on standard nonvanishing conjectures for complex L-functions.

In the sequel [Büyükboduk et al. 2018], we need a similar nonvanishing result in the case $f = g$. In this particular situation, note that we have $k_f = k_g$ and the Rankin–Selberg L-series does not possess a single critical value. In order to prove the nonvanishing of the geometric p-adic L-function associated to the symmetric square, one needs to factor the p-adic Rankin–Selberg L-function as a product of the symmetric square p-adic L-function and a Kubota–Leopoldt p-adic L-function (extending the work of Dasgupta in the p-ordinary case). This is the subject of a forthcoming work of Alessandro Arlandini.

3.9. *A partial result towards Conjecture 3.5.1.*

Theorem 3.9.1. *Suppose that all four p-adic Rankin–Selberg L-functions (3.6.1) are non-zero-divisors in $\mathcal{H}$, and that the following "big image" hypotheses hold*:

- *The representation V is absolutely irreducible.*
- *There exists an element $\tau \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}(\mu_{p^\infty}))$ such that the E-vector space $V/(\tau - 1)V$ is 1-dimensional.*

Then there exists a class $\mathrm{BF}_1 \in \mathrm{Frac}\,\mathcal{H} \otimes_\Lambda \bigwedge^2 H^1_{\mathrm{Iw}}(\mathbb{Q}(\mu_{p^\infty}), T)$ satisfying

$$\left\langle \mathcal{L}_V(\mathrm{BF}_1), v^*_{\lambda,\mu} \right\rangle = \mathrm{BF}_{\lambda,\mu,1}$$

for all choices of p-stabilisations λ, μ. In particular the "extra" antisymmetry property (3.7.1) holds.

The proof of this theorem will proceed in several steps.

Proposition 3.9.2. *If any one of the four p-adic L-functions is a non-zero-divisor and the "big image" conditions hold, then $H^1_{\mathrm{an}}(\mathbb{Q}(\mu_{p^\infty}), V)$ has rank 2 over $\mathcal{H}$, and the map*

$$H^1_{\mathrm{an}}(\mathbb{Q}(\mu_{p^\infty}), V) \to \mathcal{H}^{\oplus 4}$$

given by pairing $\mathcal{L}_V \circ \mathrm{loc}_p$ with the four basis vectors $\{v^*_{\alpha\alpha}, v^*_{\alpha\beta}, v^*_{\beta\beta}, v^*_{\beta\alpha}\}$ of $\mathbb{D}_{\mathrm{cris}}(V^*)$ (in that order) is an injection.

Proof. Since the Perrin-Riou regulator $\mathcal{L}_V$ is injective, it suffices to show that $H^1_{\mathrm{an}}(\mathbb{Q}(\mu_{p^\infty}), V)$ has rank 2 and injects into $H^1_{\mathrm{an}}(\mathbb{Q}_p(\mu_{p^\infty}), V)$ via loc_p. Note that $H^1_{\mathrm{an}}(\mathbb{Q}(\mu_{p^\infty}), V)$ is free thanks to our big image conditions, and its rank is at least 2 by Tate's Euler characteristic formula.

By symmetry we may suppose that $L(f_\alpha, g)$ is a non-zero-divisor. Choose a character χ of Γ in each Δ-isotypic component at which this p-adic L-function does not vanish, and away from the support of the torsion module $H^2_{\mathrm{Iw}}(\mathbb{Q}_p, V)$. By the explicit reciprocity law, this implies that $\mathrm{BF}_{\alpha,\alpha,1}(\chi)$ and $\mathrm{BF}_{\alpha,\beta,1}(\chi)$ are both nonzero, and moreover that their images in $H^1(\mathbb{Q}_p, V(\chi^{-1}))$ are linearly independent. By the Euler system machinery and an application of Poitou–Tate duality, precisely as in Theorem 8.2.1 and Corollary 8.3.2 of [Loeffler and Zerbes 2016b], one sees that the relaxed Selmer group $H^1_{\mathrm{relaxed}}(\mathbb{Q}, V(\chi^{-1}))$ is 2-dimensional and injects into the local cohomology at p. Since there is an injection

$$H^1_{\mathrm{an}}(\mathbb{Q}(\mu_{p^\infty}), V)/(\gamma - \chi(\gamma)) \hookrightarrow H^1_{\mathrm{relaxed}}(\mathbb{Q}, V(\chi^{-1})),$$

the result follows. □

We now assume, for the remainder of this section, that the hypotheses in the statement of Theorem 3.9.1 are satisfied. Let $\mathcal{M}$ denote the image of $H^1_{\mathrm{an}}(\mathbb{Q}(\mu_{p^\infty}), V)$ in $\mathcal{H}^{\oplus 4}$ as given by Proposition 3.9.2. Then $\mathcal{M}$ has rank 2, as we have just established. Given an element $x \in \mathcal{H}^{\oplus 4}$ and $i \in \{1, 2, 3, 4\}$, we write x_i for its i-th coordinate (understood modulo 4, so that $x_5 = x_1$). For each i, let $\mathcal{M}^{(i)} = \{x \in \mathcal{M} : x_i = 0\}$ be the kernel of the i-th coordinate projection.

We write $m^{(1)}, \dots, m^{(4)}$ for the images in $\mathcal{M}$ of the four analytic Iwasawa cohomology classes $\mathrm{BF}_1^{\alpha\alpha}$, $\mathrm{BF}_1^{\alpha\beta}$, $\mathrm{BF}_1^{\beta\beta}$, $\mathrm{BF}_1^{\beta\alpha}$ (in that order).

Proposition 3.9.3. *For each $i \in \{1, \dots, 4\}$, we have the following*:

- $m^{(i)} \in \mathcal{M}^{(i)}$;
- $(m^{(i)})_{i+1} = -(m^{(i+1)})_i$, *and this value is a non-zero-divisor*;
- $\mathcal{M}^{(i)}$ *has rank* 1;
- $\mathcal{M}^{(i)}/\langle m^{(i)} \rangle$ *is $\mathcal{H}$-torsion.*

Proof. The first two statements follow directly from the explicit reciprocity law (Theorem 3.6.5); the theorem in particular shows that the common value $(m^{(i)})_{i+1} = -(m^{(i+1)})_i$ is one of the four p-adic L-functions (3.6.1), which are non-zero-divisors by assumption.

In particular, this shows that the images of $\mathcal{M}$ under all four coordinate projections have rank ≥ 1. Since $\mathcal{M}$ has rank 2, it follows that the submodules $\mathcal{M}^{(i)}$ all have rank 1, and that each $m^{(i)}$ spans a rank 1 submodule of $\mathcal{M}^{(i)}$, so the quotient $\mathcal{M}^{(i)}/\langle m^{(i)}\rangle$ is torsion. □

Proof of Theorem 3.9.1. Let $\{u, v\}$ denote the image in $\mathcal{M}$ of a basis of $H^1_{\mathrm{Iw}}(\mathbb{Q}(\mu_{p^\infty}), T)$. Then $\{u, v\}$ is a basis of $\mathcal{M}$, and for each i, the vector

$$u_i \cdot v - v_i \cdot u$$

lies in $\mathcal{M}^{(i)}$. It is also nonzero, since u, v are linearly independent over $\mathcal{H}$. Since $\mathcal{M}^{(i)}$ has rank 1, it follows that there is a non-zero-divisor c_i in the total ring of fractions of $\mathcal{H}$ such that

$$m^{(i)} = c_i \cdot (u_i \cdot v - v_i \cdot u).$$

Substituting the definition of the c_i into the formula $(m^{(i)})_{i+1} = -(m^{(i+1)})_i$, we deduce that

$$c_i \cdot (u_i v_{i+1} - v_i u_{i+1}) = c_{i+1} \cdot (u_i v_{i+1} - v_i u_{i+1}),$$

and moreover that the common value is a non-zero-divisor. Hence $c_i = c_{i+1}$. Repeating this argument for each i, we see that the quantities c_i are all equal to some common value $c \in \mathrm{Frac}(\mathcal{H})$. Let $\tilde{u}$ and $\tilde{v}$ be the preimages of u and v in $H^1_{\mathrm{Iw}}(\mathbb{Q}(\mu_{p^\infty}), T)$. On unravelling the notation, we see that $m^{(i)}$ is equal to the image of $c \cdot \tilde{u} \wedge \tilde{v} \in \mathrm{Frac}(\mathcal{H}) \otimes_\Lambda \bigwedge^2 H^1_{\mathrm{Iw}}(\mathbb{Q}(\mu_{p^\infty}), T)$ under the map sending $x \wedge y$ to the i-th coordinate of $\left\langle \mathcal{L}_V(x)y - \mathcal{L}(y)x, v^{(i)} \right\rangle$, where $v^{(i)}$ is the i-th element of our basis $\left(v^*_{\alpha\alpha}, v^*_{\alpha\beta}, v^*_{\beta\beta}, v^*_{\beta\alpha}\right)$ of $\mathbb{D}_{\mathrm{cris}}(V)^*$. So we may take $\mathrm{BF}_1 = c \cdot \tilde{u} \wedge \tilde{v}$. □

Remark 3.9.4. We can carry out the same argument after applying the idempotent e_η, if we assume that the e_η isotypic parts of the L-functions (3.6.1) are nonzero. It suffices to have nonvanishing of any three of the four, as is clear from the proof.

More subtly, if one assumes that the symmetry property (3.7.1) for the "extra" L-functions is true, then one can prove the same theorem assuming that two of the four functions (3.6.1) and one of the "extra" L-functions is nonzero. This nonvanishing can be deduced from the explicit reciprocity law when $|k_f - k_g| \geq 3$ and appropriate η, as above.

4. Logarithmic matrix and factorisations

In the following sections of the paper, we explore some of the consequences of Conjecture 3.5.1, and show that it implies factorisations of the Beilinson–Flach elements via a matrix of logarithms.

4.1. ***Integral p-adic Hodge theory for V.*** In this section, we assume (as in the introduction) that both f and g are nonordinary at p. We also impose the following *Fontaine–Laffaille* hypothesis:

$$p > k_f + k_g + 2.$$

We describe our constructions assuming $k_f \leq k_g$ for simplicity; the case $k_f \geq k_g$ is similar.

For $h \in \{f, g\}$, we have the basis $\omega_h, \varphi(\omega_h)$ of $\mathbb{D}_{\mathrm{cris}}(R_h^*)$ with ω_h generating $\mathrm{Fil}^0\, \mathbb{D}_{\mathrm{cris}}(R_h^*)$. As worked out in [Lei et al. 2017, §3.1], the matrix of φ with respect to this basis is

$$A_h = \begin{pmatrix} 0 & -\epsilon_h(p)/p^{k_h+1} \\ 1 & a_p(h)/p^{k_h+1} \end{pmatrix}.$$

Let T be the representation $R_f^* \otimes R_g^*$. Then we consider the basis $v_1 = \omega_f \otimes \omega_g,\ v_2 = \omega_f \otimes \varphi(\omega_g)$, $v_3 = \varphi(\omega_f) \otimes \omega_g,\ v_4 = \varphi(\omega_f) \otimes \varphi(\omega_g)$ for $\mathbb{D}_{\mathrm{cris}}(T)$. We note in particular that it respects the filtration of $\mathbb{D}_{\mathrm{cris}}(T)$ in the following sense:

$$\mathrm{Fil}^i\, \mathbb{D}_{\mathrm{cris}}(T) = \begin{cases} \langle v_1, v_2, v_3, v_4 \rangle, & i \le -k_f - k_g - 2, \\ \langle v_1, v_2, v_3 \rangle, & -k_f - k_g - 1 \le i \le -k_g - 1, \\ \langle v_1, v_2 \rangle, & -k_g \le i \le -k_f - 1, \\ \langle v_1 \rangle, & -k_f \le i \le 0, \\ 0, & i \ge 1. \end{cases} \tag{4.1.1}$$

The matrix of φ with respect to this basis is

$$A = A_0 \cdot \begin{pmatrix} 1 & & & \\ & 1/p^{k_f+1} & & \\ & & 1/p^{k_g+1} & \\ & & & 1/p^{k_f+k_g+2} \end{pmatrix},$$

where A_0 is the matrix defined by

$$\begin{pmatrix} 0 & 0 & 0 & \epsilon_f(p)\epsilon_g(p) \\ 0 & 0 & -\epsilon_f(p) & -\epsilon_f(p)a_p(g) \\ 0 & -\epsilon_g(p) & 0 & -\epsilon_g(p)a_p(f) \\ 1 & a_p(g) & a_p(f) & a_p(f)a_p(g) \end{pmatrix}.$$

We introduce the following convention.

Convention 4.1.1. Let $n \ge 1$ be an integer and U an E-vector space. If $M = (m_{ij})$ is an $n \times n$ matrix defined over E and $u_1, \ldots, u_n$ are elements in U, we write

$$\begin{pmatrix} u_1 & \cdots & u_n \end{pmatrix} \cdot M$$

for the row vector of elements in U given by $\sum_{i=1}^n u_i m_{ij}$, $j = 1, \ldots, n$.

Under this convention, we have the equation

$$\begin{pmatrix} \varphi(v_1) & \varphi(v_2) & \varphi(v_3) & \varphi(v_4) \end{pmatrix} = \begin{pmatrix} v_1 & v_2 & v_3 & v_4 \end{pmatrix} \cdot A. \tag{4.1.2}$$

Recall that the Wach module $\mathbb{N}(T)$ is a free module of rank 4 over $\mathbb{A}^+_{\mathbb{Q}_p}$, equipped with a canonical isomorphism

$$\mathbb{N}(T)/\pi\mathbb{N}(T) \cong \mathbb{D}_{\mathrm{cris}}(T). \tag{4.1.3}$$

By [Berger 2004, proof of Proposition V.2.3] (see also [Lei 2017, Proposition 4.1]), our Fontaine–Laffaille hypothesis allows us to lift the basis $\{v_i\}$ of $\mathbb{D}_{\mathrm{cris}}(T)$ as an $\mathcal{O}$-module to a basis $\{n_i\}$ of $\mathbb{N}(T)$ as an $\mathbb{A}^+_{\mathbb{Q}_p}$-module, and the matrix of φ with respect to the basis $\{n_i\}$ is given by

$$P := A_0 \cdot \begin{pmatrix} \mu^{k_f+k_g+2} & & & \\ & \mu^{k_g+1}/q^{k_f+1} & & \\ & & \mu^{k_f+1}/q^{k_g+1} & \\ & & & 1/q^{k_f+k_g+2} \end{pmatrix},$$

where $\mu = p/(q-\pi^{p-1}) \in 1+\pi\mathbb{A}^+_{\mathbb{Q}_p}$. Note that P^{-1} is integral. Furthermore, similar to the Equation (4.1.2), we have

$$\begin{pmatrix}\varphi(n_1) & \varphi(n_2) & \varphi(n_3) & \varphi(n_4)\end{pmatrix} = \begin{pmatrix} n_1 & n_2 & n_3 & n_4\end{pmatrix} \cdot P. \tag{4.1.4}$$

4.2. ***The logarithmic matrix.*** There is an isomorphism

$$\mathbb{B}^+_{\mathrm{rig},\mathbb{Q}_p}\left[\tfrac{t}{\pi}\right] \otimes_{\mathbb{A}^+_{\mathbb{Q}_p}} \mathbb{N}(T) \cong \mathbb{B}^+_{\mathrm{rig},\mathbb{Q}_p}\left[\tfrac{t}{\pi}\right] \otimes_{\mathbb{Z}_p} \mathbb{D}_{\mathrm{cris}}(T) \tag{4.2.1}$$

compatible with (4.1.3) via reduction mod π. Let $M \in \mathrm{GL}_4\left(\mathbb{B}^+_{\mathrm{rig},\mathbb{Q}_p}\left[\tfrac{t}{\pi}\right]\right)$ be the matrix of this isomorphism with respect to our bases $\{v_i\}$ and $\{n_i\}$, so that

$$\begin{pmatrix} n_1 & n_2 & n_3 & n_4\end{pmatrix} = \begin{pmatrix} v_1 & v_2 & v_3 & v_4\end{pmatrix} \cdot M \tag{4.2.2}$$

under Convention 4.1.1. By [Lei 2017, Proposition 4.2], we can (and do) choose the n_i such that

$$M \equiv I_4 \mod \pi^{k_f+k_g+2}. \tag{4.2.3}$$

If we apply φ, we deduce from (4.1.2) and (4.1.4) that

$$M = A\varphi(M)P^{-1}.$$

If we repeatedly apply φ, we get

$$M = A^n\varphi^n(M)\varphi^{n-1}(P^{-1})\cdots\varphi(P^{-1})P^{-1}.$$

So, in particular,

$$M \equiv A^n\varphi^{n-1}(P^{-1})\cdots\varphi(P^{-1})P^{-1} \mod \varphi^n(\pi^{k_f+k_g+2}) \tag{4.2.4}$$

thanks to (4.2.3). We define the *logarithmic matrix* to be the 4×4 matrix over $\mathcal{H}$ given by

$$M_{\log} := \mathfrak{M}^{-1}\big((1+\pi)A\varphi(M)\big),$$

where $\mathfrak{M}$ is the Mellin transform (applied individually to each entry of the matrix $(1+\pi)A\varphi(M)$). Recall from [Lei et al. 2011, §3] that, up to a unit, the determinant of $M_{\log}$ is given by

$$\mathfrak{n}_{k_f+1}\cdot\mathfrak{n}_{k_g+1}\cdot\mathfrak{n}_{k_f+k_g+2}.$$

Furthermore, by Theorem 2.3.1 and (4.2.4), we have the congruence

$$M_{\log} \equiv A^{n+1} \cdot H_n \mod \omega_{n,k_f+k_g+2}, \tag{4.2.5}$$

where $H_n = \mathfrak{M}^{-1}(\varphi^n(P^{-1}) \cdots \varphi(P^{-1}))$.

Lemma 4.2.1. *The adjugate matrix* $\mathrm{adj}(M_{\log}) = \det(M_{\log})M_{\log}^{-1}$ *is divisible by* $\mathfrak{n}_{k_f+1}\mathfrak{n}_{k_g+1}$.

Proof. By (4.2.5), we have

$$M_{\log} \equiv A^{n+1} \cdot H_n \mod \Phi_{n,k_f+k_g+2}.$$

So, it is enough to show that $\mathrm{adj}(H_n)$ is divisible by $\Phi_{n,k_f+1}\Phi_{n,k_g+1}$ for all n. Recall that,

$$\mathfrak{M}(H_n) = \varphi^n(P^{-1}) \cdots \varphi(P^{-1}).$$

From the construction of P, the last three rows of P^{-1} are divisible by q^{k_f+1}, q^{k_g+1} and $q^{k_f+k_g+2}$ respectively. Therefore, the last three rows of $\mathfrak{M}(H_n)$ are divisible by $\varphi^n(q^{k_f+1})$, $\varphi^n(q^{k_g+1})$ and $\varphi^n(q^{k_f+k_g+2})$ respectively. Theorem 2.3.1 then tells us that the last three rows of H_n are divisible by Φ_{n,k_f+1}, Φ_{n,k_g+1} and Φ_{n,k_f+k_g+2}. Hence, when we take adjugate, every entry will be divisible by $\Phi_{n,k_f+1}\Phi_{n,k_g+1}$ as required. □

Let $\{v_{\lambda,\mu}\}_{\lambda,\mu\in\{\alpha,\beta\}}$ be the eigenvector basis of $\mathbb{D}_{\mathrm{cris}}(V)$ as given in Section 3.5. The matrix of φ with respect to this basis is

$$D := \begin{pmatrix} \frac{1}{\alpha_f\alpha_g} & & & \\ & \frac{1}{\alpha_f\beta_g} & & \\ & & \frac{1}{\beta_f\alpha_g} & \\ & & & \frac{1}{\beta_f\beta_g} \end{pmatrix}.$$

Recall that we defined $v_{\lambda\mu} = v^*_{f,\lambda} \otimes v^*_{g,\mu}$, where $v_{h,\alpha}$ and $v_{h,\beta}$ are eigenvectors in $\mathbb{D}_{\mathrm{cris}}(V_h)$ with $v_{h,\alpha} = v_{h,\beta} \mod \mathrm{Fil}^1$ for $h \in \{f, g\}$. Since $\langle v^*_{h,\alpha} + v^*_{h,\beta}, v_{h,\alpha} - v_{h,\beta}\rangle = 0$ by duality, we have $v^*_{h,\alpha} + v^*_{h,\beta} = 0 \mod \mathrm{Fil}^0$. After multiplying ω_h by a scalar if necessary, we may choose

$$v^*_{h,\alpha} = \alpha_h(\omega_h - \beta_h\varphi(\omega_h)), \quad \text{and} \quad v^*_{h,\beta} = -\beta_h(\omega_h - \alpha_h\varphi(\omega_h)).$$

We let Q be the change-of-basis matrix from the basis $\{v_i\}$ to this eigenvector basis, so that $D = Q^{-1}AQ$. Explicitly, we have

$$Q = \begin{pmatrix} \alpha_f\alpha_g & -\alpha_f\beta_g & -\beta_f\alpha_g & \beta_f\beta_g \\ -\alpha_f\alpha_g\beta_g & \alpha_f\alpha_g\beta_g & -\alpha_g\beta_f\beta_g & \alpha_g\beta_f\beta_g \\ -\alpha_f\alpha_g\beta_f & \alpha_f\beta_f\beta_g & -\alpha_f\alpha_g\beta_f & \alpha_f\beta_f\beta_g \\ \alpha_f\alpha_g\beta_f\beta_g & -\alpha_f\alpha_g\beta_f\beta_g & -\alpha_f\alpha_g\beta_f\beta_g & \alpha_f\alpha_g\beta_f\beta_g \end{pmatrix}.$$

Using this matrix, we may rewrite (4.2.5) as

$$Q^{-1}M_{\log} \equiv D^{n+1}Q^{-1}H_n \mod \omega_{n,k_f+k_g+2}. \tag{4.2.6}$$

Lemma 4.2.2. *The entries in the first column of $Q^{-1}M_{\log}$ are all $O(\log_p^{v_p(\alpha_f\alpha_g)})$, and similarly for the other three columns.*

Proof. Recall that H_n is a matrix defined over Λ. The congruence relation (4.2.6) tells us that the entries of the first column of $Q^{-1}M_{\log}$ modulo ω_{n,k_f+k_g+2} have denominator $O(p^{v_p(\alpha_f\alpha_g)})$. Hence, our result follows from [Büyükboduk and Lei 2016, Lemma 2.2], which is a slight generalisation of [Perrin-Riou 1994, §1.2.1]. □

4.3. ***Characterising the image.*** Here we prove an important linear-algebra result describing the Λ-submodule of $\mathcal{H}^{\oplus 4}$ generated by the logarithmic matrix; we shall see that it consists exactly of those elements which "look like" they are in the image of the Perrin-Riou regulator map.

Proposition 4.3.1. *Let $F_{\lambda,\mu} \in \mathcal{H}$, $\lambda, \mu \in \{\alpha, \beta\}$, be four functions. Suppose that, for some integer $j \in \{0, \dots, k_f + k_g + 1\}$ and some Dirichlet character θ of conductor p^n with $n > 1$, we have*

$$\sum_{\lambda,\mu} (\lambda\mu)^n F_{\lambda,\mu}(\chi^j\theta) v_{\lambda,\mu} \in \mathbb{Q}_{p,n} \otimes \mathrm{Fil}^{-j}\, \mathbb{D}_{\mathrm{cris}}(T).$$

Then,

$$\frac{\mathrm{adj}(Q^{-1}M_{\log})}{\mathfrak{n}_{k_f+1}\mathfrak{n}_{k_g+1}} \cdot \begin{pmatrix} F_{\alpha,\alpha} \\ F_{\alpha,\beta} \\ F_{\beta,\alpha} \\ F_{\beta,\beta} \end{pmatrix} (\chi^j\theta) = 0.$$

Proof. When we evaluate an element in Λ at $\chi^j\theta$, the result only depends on the given element modulo $\mathrm{Tw}^{-j}\Phi_{n-1}(X)$. By (4.2.5),

$$\mathrm{adj}(Q^{-1}M_{\log}) \equiv \mathrm{adj}(D^n Q^{-1} H_{n-1}) \equiv \frac{\det(D^n)}{\det(Q)} \mathrm{adj}(H_{n-1}) Q D^{-n} \mod \mathrm{Tw}^{-j}\Phi_{n-1}.$$

We recall from the proof of Lemma 4.2.1 that the last three rows of H_{n-1} are divisible by Φ_{n-1,k_f+1}, Φ_{n-1,k_g+1} and Φ_{n-1,k_f+k_g+2} respectively. So, after dividing by $\mathfrak{n}_{k_f+1}\mathfrak{n}_{k_g+1}$, the first column of $\mathrm{adj}(H_{n-1})$ is divisible by Φ_{n-1,k_f+k_g+2}, the second column is divisible by $\Phi_{n-1,k_f+k_g+2}/\Phi_{n-1,k_f+1}$, whereas the third one is divisible by $\Phi_{n-1,k_f+k_g+2}/\Phi_{n-1,k_g+1}$. In particular, when evaluated at a character of the form $\chi^j\theta$, we have

$$\frac{\mathrm{adj}(H_{n-1})}{\mathfrak{n}_{k_f+1}\mathfrak{n}_{k_g+1}}(\chi^j\theta) = \begin{pmatrix} 0 & 0 & 0 & * \\ 0 & 0 & 0 & * \\ 0 & 0 & 0 & * \\ 0 & 0 & 0 & * \end{pmatrix}$$

if $k_g + 1 \le j \le k_f + k_g + 1$. When j is in this range, our assumption on $F_{\lambda,\mu}$ tells us that

$$QD^{-n} \begin{pmatrix} F_{\alpha,\alpha} \\ F_{\alpha,\beta} \\ F_{\beta,\alpha} \\ F_{\beta,\beta} \end{pmatrix} (\chi^j\theta) = \begin{pmatrix} * \\ * \\ * \\ 0 \end{pmatrix}$$

thanks to the description of the filtration in (4.1.1). The result then follows from multiplying the two equations above.

For the other cases, we have

$$\frac{\mathrm{adj}(H_{n-1})}{\mathfrak{n}_{k_f+1}\mathfrak{n}_{k_g+1}}(\chi^j\theta) = \begin{pmatrix} 0 & 0 & * & * \\ 0 & 0 & * & * \\ 0 & 0 & * & * \\ 0 & 0 & * & * \end{pmatrix}, \quad QD^{-n}\begin{pmatrix} F_{\alpha,\alpha} \\ F_{\alpha,\beta} \\ F_{\beta,\alpha} \\ F_{\beta,\beta} \end{pmatrix}(\chi^j\theta) = \begin{pmatrix} * \\ * \\ 0 \\ 0 \end{pmatrix}$$

if $k_f - 1 \leq j \leq k_g$ and

$$\frac{\mathrm{adj}(H_{n-1})}{\mathfrak{n}_{k_f+1}\mathfrak{n}_{k_g+1}}(\chi^j\theta) = \begin{pmatrix} 0 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \end{pmatrix}, \quad QD^{-n}\begin{pmatrix} F_{\alpha,\alpha} \\ F_{\alpha,\beta} \\ F_{\beta,\alpha} \\ F_{\beta,\beta} \end{pmatrix}(\chi^j\theta) = \begin{pmatrix} * \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

if $0 \leq j \leq k_f$, so we are done. □

Theorem 4.3.2. *Let $F_{\lambda,\mu} \in \mathcal{H}$, $\lambda, \mu \in \{\alpha, \beta\}$ be four functions such that for all integers $0 \leq j \leq k_f + k_g + 1$ and all Dirichlet characters θ of conductor p^n with $n > 1$,*

$$\sum_{\lambda,\mu}(\lambda\mu)^n F_{\lambda,\mu}(\chi^j\theta)v_{\lambda,\mu} \in \mathbb{Q}_{p,n} \otimes \mathrm{Fil}^{-j}\,\mathbb{D}_{\mathrm{cris}}(T).$$

Then,

$$\begin{pmatrix} F_{\alpha,\alpha} \\ F_{\alpha,\beta} \\ F_{\beta,\alpha} \\ F_{\beta,\beta} \end{pmatrix} = Q^{-1}M_{\log} \cdot \begin{pmatrix} F_{\#,\#} \\ F_{\#,\flat} \\ F_{\flat,\#} \\ F_{\flat,\flat} \end{pmatrix}$$

for some $F_{\bullet,\circ} \in \mathcal{H}$.

Furthermore, if $F_{\lambda,\mu} = O(\log_p^{v_p(\lambda_f\mu_g)})$ for all four choices of λ and μ, then $F_{\bullet,\circ} = O(1)$ for all $\bullet$ and $\circ$.

Proof. Proposition 4.3.1 tells us that

$$\frac{\mathrm{adj}(Q^{-1}M_{\log})}{\mathfrak{n}_{k_f+1}\mathfrak{n}_{k_g+1}} \cdot \begin{pmatrix} F_{\alpha,\alpha} \\ F_{\alpha,\beta} \\ F_{\beta,\alpha} \\ F_{\beta,\beta} \end{pmatrix} \in \mathfrak{n}_{k_f+k_g+2}\mathcal{H}^{\oplus 4}.$$

But since the determinant of $Q^{-1}M_{\log}$ is up to a unit $\mathfrak{n}_{k_f+1}\mathfrak{n}_{k_g+1}\mathfrak{n}_{k_f+k_g+2}$,

$$\frac{\mathrm{adj}(Q^{-1}M_{\log})}{\mathfrak{n}_{k_f+1}\mathfrak{n}_{k_g+1}\mathfrak{n}_{k_f+k_g+2}}$$

is (again up to a unit) $(Q^{-1}M_{\log})^{-1}$, hence the decomposition as claimed.

If furthermore $F_{\lambda,\mu} = O(\log_p^{v_p(\lambda_f\mu_g)})$, then Lemma 4.2.2 tells us that all entries in the product

$$\frac{\mathrm{adj}(Q^{-1}M_{\log})}{\mathfrak{n}_{k_f+1}\mathfrak{n}_{k_g+1}} \cdot \begin{pmatrix} F_{\alpha,\alpha} \\ F_{\alpha,\beta} \\ F_{\beta,\alpha} \\ F_{\beta,\beta} \end{pmatrix}$$

are $O(\log_p^{k_f+k_g+2})$. This says that the quotient

$$\frac{\mathrm{adj}(Q^{-1}M_{\log})}{\mathfrak{n}_{k_f+1}\mathfrak{n}_{k_g+1}\mathfrak{n}_{k_f+k_g+2}} \cdot \begin{pmatrix} F_{\alpha,\alpha} \\ F_{\alpha,\beta} \\ F_{\beta,\alpha} \\ F_{\beta,\beta} \end{pmatrix}$$

is in $O(1)$ and we are done. □

Remark 4.3.3. Note that the condition on $F_{\alpha\beta}$ is automatically satisfied if the $F_{\alpha\beta}$ are the components in our eigenvector basis of an element of the Perrin-Riou regulator map, since the regulator interpolates the Bloch–Kato dual exponential for $j \geq 0$, and the dual exponential map for $T(-j)$ factors through $\mathrm{Fil}^{-j}\mathbb{D}_{\mathrm{cris}}$. The above result should be viewed as a sort of converse to this statement, showing that these vanishing conditions force a factorisation via the matrix of logarithms, of the same form as the factorisation established for the Perrin-Riou regulator in [Lei et al. 2010].

5. Equivariant Perrin-Riou maps and (#, ♭)-splitting

5.1. ***Perrin-Riou maps and signed Coleman map.*** Let f and g be two modular forms as in the previous section. We shall write $T = R_f^* \otimes R_g^*$ as before. Let $F/\mathbb{Q}_p$ be a finite unramified extension. We write $\mathbb{N}_F(T)$ and $\mathbb{D}_{\mathrm{cris}}(F, T)$ for the Wach module and Dieudonné module of T over F. We have already fixed bases $\{n_i\}$ and $\{v_i\}$ of $\mathbb{N}_{\mathbb{Q}_p}(T)$ and $\mathbb{D}_{\mathrm{cris}}(\mathbb{Q}_p, T)$ respectively. Given that

$$\mathbb{N}_F(T) = \mathcal{O}_F \otimes_{\mathbb{Z}_p} \mathbb{N}_{\mathbb{Q}_p}(T), \quad \mathbb{D}_{\mathrm{cris}}(F, T) = \mathcal{O}_F \otimes_{\mathbb{Z}_p} \mathbb{D}_{\mathrm{cris}}(\mathbb{Q}_p, T),$$

we may extend the bases we have chosen to $\mathbb{N}_F(T)$ and $\mathbb{D}_{\mathrm{cris}}(F, T)$ naturally.

We recall from (4.2.2) that the change of basis matrix M between the two bases above results in a logarithmic matrix $M_{\log}$. Furthermore, given that $M \equiv I_4 \mod \pi^2$ by (4.2.4), [Lei et al. 2017, Theorem 2.5] says that $\{(1+\pi)\varphi(n_i)\}$ is a Λ-basis of $\varphi^*(\mathbb{N}(T))$. We recall from [Lei et al. 2010, Remark 3.4] that $(1-\varphi)\mathbb{N}_F(T)^{\psi=1} \subset (\varphi^*\mathbb{N}_F(T))^{\psi=0}$. This allows us to define four Coleman maps $\mathrm{Col}_{F,\bullet,\circ} : \mathbb{N}_F(T)^{\psi=1} \to \mathcal{O}_F \otimes \Lambda$ via the relation

$$(1-\varphi)z = \begin{pmatrix} v_1 & v_2 & v_3 & v_4 \end{pmatrix} \cdot M_{\log} \cdot \begin{pmatrix} \mathrm{Col}_{F,\#,\#}(z) \\ \mathrm{Col}_{F,\#,\flat}(z) \\ \mathrm{Col}_{F,\flat,\#}(z) \\ \mathrm{Col}_{F,\flat,\flat}(z) \end{pmatrix}$$

for $z \in \mathbb{N}_F(T)^{\psi=1}$ (see [Lei et al. 2011, §3] for details).

A result of Berger [2003, Theorem A.3] tells us that there is an isomorphism

$$h^1_{F,T} : \mathbb{N}_F(T)^{\psi=1} \to H^1_{\mathrm{Iw}}(F, T).$$

We shall abuse notation and denote $\mathrm{Col}_{F,\bullet,\circ} \circ (h^1_{F,T})^{-1}$ by simply $\mathrm{Col}_{F,\bullet,\circ}$ for $\bullet, \circ \in \{\#, \flat\}$. The Perrin-Riou regulator map

$$\mathcal{L}_{T,F} : H^1_{\mathrm{Iw}}(F, T) \to \mathcal{H} \otimes \mathbb{D}_{\mathrm{cris}}(F, T)$$

is given by

$$(\mathfrak{M}^{-1} \otimes 1) \circ (1 - \varphi) \circ (h^1_{F,T})^{-1}.$$

Hence, we have the decomposition

$$\mathcal{L}_{T,F}(z) = \begin{pmatrix} v_1 & v_2 & v_3 & v_4 \end{pmatrix} \cdot M_{\log} \cdot \begin{pmatrix} \mathrm{Col}_{F,\#,\#}(z) \\ \mathrm{Col}_{F,\#,\flat}(z) \\ \mathrm{Col}_{F,\flat,\#}(z) \\ \mathrm{Col}_{F,\flat,\flat}(z) \end{pmatrix}. \tag{5.1.1}$$

Exactly as in (3.3.1), for $m \geq 1$ we can combine the maps $\mathrm{Col}_{\mathbb{Q}(\mu_m)_v,\bullet,\circ}$ for primes $v \mid p$ of $\mathbb{Q}(\mu_m)$, and the map ν_m, to obtain maps

$$\mathrm{Col}_{m,\bullet,\circ} : H^1_{\mathrm{Iw}}(\mathbb{Q}(\mu_{mp^\infty}) \otimes \mathbb{Q}_p, T) \to \Lambda_m.$$

5.2. ***Signed Selmer groups.*** Let Λ^ι be the free rank 1 Λ-module on which $G_{\mathbb{Q}}$ acts via the inverse of the canonical character $G_{\mathbb{Q}} \twoheadrightarrow \Gamma \hookrightarrow \Lambda^\times$. We write $\mathbb{T} := T \otimes \Lambda^\iota$, and we define the (compact) *signed Selmer group* $H^1_{\mathcal{F}_{\bullet,\circ}}(\mathbb{Q}(\mu_m), \mathbb{T})$ by setting

$$H^1_{\mathcal{F}_{\bullet,\circ}}(\mathbb{Q}(\mu_m), \mathbb{T}) := \ker\left(H^1(\mathbb{Q}(\mu_m), \mathbb{T}) \to \prod_{v \mid p} \frac{H^1(\mathbb{Q}(\mu_m)_v, \mathbb{T})}{\ker(\mathrm{Col}_{\bullet,\circ,\mathbb{Q}(\mu_m)_v})} \right).$$

We next define discrete signed Selmer groups for the dual Galois representation $T^\vee(1)$. Let $F/\mathbb{Q}_p$ be a finite unramified extension. By Tate duality, there is a perfect pairing

$$H^1_{\mathrm{Iw}}(F, T) \times H^1(F(\mu_{p^\infty}), T^\vee(1)) \to \mathbb{Q}_p/\mathbb{Z}_p.$$

For $\bullet, \circ \in \{\#, \flat\}$, we define

$$H^1_{\bullet,\circ}(F(\mu_{p^\infty}), T^\vee(1)) \subset H^1(F(\mu_{p^\infty}), T^\vee(1))$$

to be the orthogonal complement of $\ker(\mathrm{Col}_{\bullet,\circ,F})$.

Definition 5.2.1. The discrete signed Selmer group $\mathrm{Sel}_{\bullet,\circ}(T^\vee(1)/\mathbb{Q}(\mu_{mp^\infty}))$ is the kernel of the restriction map

$$H^1(\mathbb{Q}(\mu_{mp^\infty}), T^\vee(1)) \to \prod_{v \mid p} \frac{H^1(\mathbb{Q}(\mu_{mp^\infty})_v, T^\vee(1))}{H^1_{\bullet,\circ}(\mathbb{Q}(\mu_{mp^\infty})_v, T^\vee(1))} \times \prod_{v \nmid p} \frac{H^1(\mathbb{Q}(\mu_{mp^\infty})_v, T^\vee(1))}{H^1_f(\mathbb{Q}(\mu_{mp^\infty})_v, T^\vee(1))},$$

where v runs through all primes of $\mathbb{Q}(\mu_{mp^\infty})$.

5.3. ***(#, ♭)-splitting and rank-2 Euler systems.*** Our goal in this section is to formulate a weaker alternative to Conjecture 3.5.1, which we are able to verify in many cases of interest (in [Büyükboduk and Lei 2016; Büyükboduk et al. 2018]), allowing us to make full use of the Euler system machinery in these scenarios.

Conjecture 5.3.1. *There exists a nonzero $r_0 \in \mathbb{Z}$, and a collection of elements $\mathrm{BF}_{\bullet,\circ,m} \in H^1_{\mathcal{F}_{\bullet,\circ}}(\mathbb{Q}(\mu_m), \mathbb{T})$ for each $m \in \mathcal{N}(\mathcal{P})$ and each choice of $\bullet, \circ \in \{\#, \flat\}$, such that*

$$r_0 \cdot \begin{pmatrix} \mathrm{BF}_{\alpha,\alpha,m} \\ \mathrm{BF}_{\alpha,\beta,m} \\ \mathrm{BF}_{\beta,\alpha,m} \\ \mathrm{BF}_{\beta,\beta,m} \end{pmatrix} = Q^{-1} M_{\log} \cdot \begin{pmatrix} \mathrm{BF}_{\#,\#,m} \\ \mathrm{BF}_{\#,\flat,m} \\ \mathrm{BF}_{\flat,\#,m} \\ \mathrm{BF}_{\flat,\flat,m} \end{pmatrix}. \tag{5.3.1}$$

Note that the $\mathrm{BF}_{\bullet,\circ,m}$ are uniquely determined if they exist, since the determinant of $Q^{-1}M_{\log}$ is a non-zero-divisor in $\mathcal{H}$.

Proposition 5.3.2. *If Conjecture 3.5.1 holds, then Conjecture 5.3.1 holds (and we may take $r_0 = 1$).*

Proof. Suppose that Conjecture 3.5.1 is true. For each $m \in \mathcal{N}(\mathcal{P})$ and each $\bullet, \circ \in \{\#, \flat\}$, we can regard $\mathrm{Col}_{m,\bullet,\circ} \circ \mathrm{loc}_p$ as a map

$$\bigwedge^2 H^1_{\mathrm{Iw}}(\mathbb{Q}(\mu_{mp^\infty}), T) \to H^1_{\mathrm{Iw}}(\mathbb{Q}(\mu_{mp^\infty}), T).$$

Let us set $\mathrm{BF}_{\bullet,\circ,m} := \mathrm{Col}_{\bullet,\circ,m}(\mathrm{BF}_m) \in H^1_{\mathrm{Iw}}(\mathbb{Q}(\mu_m), T)$, where BF_m is the element of Conjecture 3.5.1. Then it is clear that $\mathrm{BF}_{\bullet,\circ,m} \in H^1_{\mathcal{F}_{\bullet,\circ}}(\mathbb{Q}(\mu_m), \mathbb{T})$; and the formula (5.1.1) relating the Perrin-Riou regulator $\mathcal{L}_{V,m}$ to the Coleman maps implies that we have

$$\begin{pmatrix} \langle \mathcal{L}_{m,V}(\mathrm{BF}_m), v^*_{\alpha,\alpha} \rangle \\ \langle \mathcal{L}_{m,V}(\mathrm{BF}_m), v^*_{\alpha,\beta} \rangle \\ \langle \mathcal{L}_{m,V}(\mathrm{BF}_m), v^*_{\beta,\alpha} \rangle \\ \langle \mathcal{L}_{m,V}(\mathrm{BF}_m), v^*_{\beta,\beta} \rangle \end{pmatrix} = Q^{-1} M_{\log} \cdot \begin{pmatrix} \mathrm{BF}_{\#,\#,m} \\ \mathrm{BF}_{\#,\flat,m} \\ \mathrm{BF}_{\flat,\#,m} \\ \mathrm{BF}_{\flat,\flat,m} \end{pmatrix},$$

where $v^*_{\alpha\alpha}, \ldots, v^*_{\beta\beta}$ is the eigenvector basis of $\mathbb{D}_{\mathrm{cris}}(V^*)$. By the defining property of BF_m, we have $\langle \mathcal{L}_{V,m}(\mathrm{BF}_m), v^*_{\lambda,\mu} \rangle = \mathrm{BF}_{\lambda,\mu,m}$ for each λ, μ, which gives the required factorisation. □

Consider the following antisymmetry condition:

(**A–Sym**) For all possible choices of the symbols $\triangle, \square, \bullet, \circ \in \{\#, \flat\}$ and every $m \in \mathcal{N}(\mathcal{P})$ we have

$$\mathrm{Col}_{m,\triangle,\square} \circ \mathrm{res}_p(\mathrm{BF}_{\bullet,\circ,m}) = -\mathrm{Col}_{m,\bullet,\circ} \circ \mathrm{res}_p(\mathrm{BF}_{\triangle,\square,m}).$$

Remark 5.3.3. Assume that Conjecture 5.3.1 and the hypothesis (**A–Sym**) hold true. Then for each choice of $\bullet, \circ \in \{\#, \flat\}$, the collection $\{\mathrm{BF}_{\bullet,\circ,m}\}_m$ is a (rank 1) *locally restricted Euler system* in the sense of [Büyükboduk and Lei 2015, Appendix A], since each collection of p-stabilised Beilinson–Flach classes $\{\mathrm{BF}_{\lambda,\mu,m}\}_m$ (for $\lambda, \mu \in \{\alpha, \beta\}$) verifies the Euler system distribution relations as m varies.

See Section 5.4, where we partially verify Conjecture 5.3.1 and Proposition 5.3.4, where we give a sufficient condition for the validity of (**A–Sym**). In the sequel [Büyükboduk et al. 2018], we prove an appropriate variant of this conjecture for the twist $\mathrm{Sym}^2 f \otimes \chi$ of the symmetric square motive with a Dirichlet character.

Finally, we note that a factorisation similar to (5.3.1) is proved in [Büyükboduk and Lei 2016] unconditionally when the newform g is taken to be a p-ordinary CM form.

Proposition 5.3.4. *Suppose one of the following conditions*:

(i) *Conjecture 3.5.1 holds*;

(ii) *Conjecture 5.3.1 holds and the hypotheses of Theorem 3.9.1 are satisfied for Rankin–Selberg convolutions $f \otimes g \otimes \eta$, where η runs through characters of* $\mathrm{Gal}(\mathbb{Q}(m)/\mathbb{Q})$ *with $m \in \mathcal{N}(\mathcal{P})$.*

Then **(A–Sym)** *holds true.*

Proof. If (i) holds, then the elements $\mathrm{BF}_{\bullet,\circ,m}$ must arise as the images of BF_1 under the Coleman maps, as in the preceding proposition (by the uniqueness of the factorisation (5.3.1)); the above symmetry property is then obvious. If we assume (ii) holds, then we may carry out exactly the same argument on passing to η-isotypic components (where η runs through characters of $\mathrm{Gal}(\mathbb{Q}(m)/\mathbb{Q})$) and after extending scalars to $\mathrm{Frac}\,\mathcal{H}$. □

5.4. ***Partial* (#, ♭)*-splitting of Beilinson–Flach classes.*** Here we give evidence towards Conjecture 5.3.1 by proving a partial (#, ♭)-factorisation of Beilinson–Flach classes.

Let $m \geq 1$ be an integer coprime to p. We write

$$\mathrm{BF}_{\lambda,\mu,m} \in H^1\big(\mathbb{Q}(m),\, R_f^* \otimes R_g^* \otimes \mathcal{H}^\iota\big)$$

for the Beilinson–Flach element at tame level m associated to the p-stabilisations f^λ and g^μ.

Theorem 5.4.1. *Let $h = \max(k_f, k_g)$. Then there exist $\widetilde{\mathrm{BF}}_{\bullet,\circ,m} \in H^1_{\mathrm{Iw}}\big(\mathbb{Q}(m),\, R_f^* \otimes R_g^* \otimes \mathcal{H}^\iota\big)$, for each $\bullet, \circ \in \{\#, \flat\}^2$, such that*

$$\frac{\mathfrak{n}_{k_f+k_g+2}}{\mathfrak{n}_{h+1}} \begin{pmatrix} \mathrm{BF}_{\alpha,\alpha,m} \\ \mathrm{BF}_{\alpha,\beta,m} \\ \mathrm{BF}_{\beta,\alpha,m} \\ \mathrm{BF}_{\beta,\beta,m} \end{pmatrix} = Q^{-1} M_{\log} \cdot \begin{pmatrix} \widetilde{\mathrm{BF}}_{\#,\#,m} \\ \widetilde{\mathrm{BF}}_{\#,\flat,m} \\ \widetilde{\mathrm{BF}}_{\flat,\#,m} \\ \widetilde{\mathrm{BF}}_{\flat,\flat,m} \end{pmatrix}.$$

Proof. We fix a basis $\{z_i\}$ of $H^1\big(\mathbb{Q}(m),\, R_f^* \otimes R_g^* \otimes \Lambda^\iota\big)$ and write $\mathrm{BF}_{\lambda,\mu,m} = \sum F_{\lambda,\mu,i} z_i$ for some $F_{\lambda,\mu,i} \in \mathcal{H}$. For a fixed i, the coefficients $F_{\lambda,\mu,i}$ satisfy the conditions given in Proposition 4.3.1 for $0 \leq j \leq h$ and all θ of conductor $p^n > 1$, thanks to Lemma 3.4.3 (using case (i) of the lemma for $0 \leq j \leq \min(k_f, k_g)$, and case (ii) for $\min(k_f, k_g) < j \leq h$). Therefore,

$$\frac{\mathrm{adj}(Q^{-1} M_{\log})}{\mathfrak{n}_{k_f+1}\mathfrak{n}_{k_g+1}} \cdot \begin{pmatrix} F_{\alpha,\alpha,m} \\ F_{\alpha,\beta,m} \\ F_{\beta,\alpha,m} \\ F_{\beta,\beta,m} \end{pmatrix} \in \mathfrak{n}_{h+1} \mathcal{H}^{\oplus 4}.$$

Hence the result on multiplying $\mathfrak{n}_{k_f+k_g+2}/\mathfrak{n}_{h+1}$ on both sides and the fact that

$$\frac{\mathrm{adj}(Q^{-1} M_{\log})}{\mathfrak{n}_{k_f+1}\mathfrak{n}_{k_g+1}\mathfrak{n}_{k_f+k_g+2}}$$

is up to a unit $(Q^{-1}M_{\log})^{-1}$. □

Remark 5.4.2. Clearly, if one could show that the coefficients $F_{\lambda,\mu,i}$ also satisfied the conditions of Proposition 4.3.1 in the "antigeometric" range $\max(k_f, k_g)+1 \leq j \leq k_f+k_g+1$, then the same argument as above would prove the full strength of Conjecture 5.3.1. However, we have not been able to prove this.

6. Signed main conjectures

We shall start this section with the definition of quadruply signed Selmer groups associated to the Rankin–Selberg product $f \otimes g$. We expect that the quadruply signed Selmer groups approximate (in an appropriate sense) the Bloch–Kato Selmer groups over finite layers of the cyclotomic tower. Unfortunately, we are unable to present a justification of this expectation.

We shall formulate our quadruply signed Iwasawa main conjecture that relates the quadruply signed Selmer groups to quadruply signed p-adic L-functions (which we also define in this section). Assuming the validity of Conjecture 5.3.1 (signed-splitting for Beilinson–Flach elements), we shall prove (under mild hypotheses) a divisibility towards the quadruply signed Iwasawa main conjecture.

6.1. *Quadruply signed Selmer groups and p-adic L-functions.*

Definition 6.1.1. Let $\mathcal{S}$ denote the set of unordered pairs $\{(\triangle, \square), (\bullet, \circ)\}$ of ordered pairs, where each of $\triangle, \square, \bullet, \circ$ is one of the symbols $\{\#, \flat\}$, and $(\triangle, \square) \neq (\bullet, \circ)$.

Note that $\mathcal{S}$ has 6 elements. We shall define a Selmer group, and formulate a main conjecture, for each $\mathfrak{S} \in \mathcal{S}$.

Definition 6.1.2. Let $\mathfrak{S} = \{(\triangle, \square), (\bullet, \circ)\} \in \mathcal{S}$. We define the following objects:

- A compact Selmer group $H^1_{\mathcal{F}_{\mathfrak{S}}}(\mathbb{Q}, \mathbb{T})$, given by
$$H^1_{\mathcal{F}_{\mathfrak{S}}}(\mathbb{Q}, \mathbb{T}) := \ker\left(H^1(\mathbb{Q}, \mathbb{T}) \to \frac{H^1(\mathbb{Q}_p, \mathbb{T})}{\ker(\mathrm{Col}_{\triangle,\square,\mathbb{Q}_p}) \cap \ker(\mathrm{Col}_{\bullet,\circ,\mathbb{Q}_p})}\right).$$
- A discrete Selmer group $\mathrm{Sel}_{\mathfrak{S}}(T^\vee(1)/\mathbb{Q}(\mu_{p^\infty}))$, given by the kernel of the restriction map
$$H^1(\mathbb{Q}(\mu_{p^\infty}), T^\vee(1)) \to \prod_{v \mid p} \frac{H^1(\mathbb{Q}(\mu_{p^\infty})_v, T^\vee(1))}{H^1_{\mathfrak{S}}(\mathbb{Q}(\mu_{p^\infty})_v, T^\vee(1))} \times \prod_{v \nmid p} \frac{H^1(\mathbb{Q}(\mu_{p^\infty})_v, T^\vee(1))}{H^1_f(\mathbb{Q}(\mu_{p^\infty})_v, T^\vee(1))},$$
where v runs through all primes of $\mathbb{Q}(\mu_{p^\infty})$, and for $v \mid p$ the local condition $H^1_{\mathfrak{S}}(\mathbb{Q}(\mu_{p^\infty})_v, T^\vee(1))$ is the orthogonal complement of $\ker(\mathrm{Col}_{\triangle,\square,\mathbb{Q}_p}) \cap \ker(\mathrm{Col}_{\bullet,\circ,\mathbb{Q}_p})$ under the local Tate pairing.
- Assuming the hypotheses of Proposition 5.3.4, we define a *quadruply signed p-adic L-function* by
$$\mathfrak{L}_{\mathfrak{S}} := \mathrm{Col}_{\triangle,\square,\mathbb{Q}_p} \circ \mathrm{res}_p(\mathrm{BF}_{\bullet,\circ,1}) \in \Lambda.$$

Remark 6.1.3. The element $\mathfrak{L}_{\mathfrak{S}}$ is only well-defined up to sign, since interchanging the role of $(\triangle, \square)$ and $(\bullet, \circ)$ has the effect of multiplying the p-adic L-function by -1 (this is the content of Proposition 5.3.4).

However, we shall only be interested in the ideal generated by $\mathfrak{L}_{\mathfrak{S}}$, so the ambiguity of signs is no problem for us.

Remark 6.1.4. We conjecture below an explicit relation between the quadruply signed Selmer group and the quadruply signed p-adic L-function, and offer some partial results towards its validity. For motivational purposes, we shall provide here one philosophical reason why quadruply signed Selmer groups are the correct choice (over doubly signed Selmer groups).

Since rank $T^- =$ rank $T^+ = 2$, one may deduce using Poitou–Tate global duality (as utilised in the proof of Theorem 5.2.15 and Lemma 5.3.16 of [Mazur and Rubin 2004]) that

$$\operatorname{rank}_\Lambda H^1_{\mathcal{F}_?}(\mathbb{Q}, \mathbb{T}) - \operatorname{rank}_\Lambda \operatorname{Sel}_?\big(T^\vee(1)/\mathbb{Q}(\mu_{p^\infty})\big) = \begin{cases} 1 & \text{if } ? = (\bullet, \circ) \in \{\#, \flat\}^2, \\ 0 & \text{if } ? = \mathfrak{S} \in \mathcal{S}, \end{cases}$$

and one expects, in the spirit of weak Leopoldt conjecture, that $\operatorname{rank}_\Lambda H^1_{\mathcal{F}_?}(\mathbb{Q}, \mathbb{T})$ should be as small as possible subject to these conditions. Moreover, in line with Bloch–Kato conjectures, one would also expect that $\operatorname{rank}_\Lambda H^1_{\mathcal{F}_?}(\mathbb{Q}, \mathbb{T})$ is given as the generic order of vanishing of (appropriate linear combinations of) L-values associated to the motives $M(f) \otimes M(g) \otimes \chi$, where χ ranges among Dirichlet characters of p-power conductor. In the critical range, note that the generic order of vanishing of these L-values is zero and this should also be the case for at least one of the said linear combinations. This tells us that the corresponding Selmer group ought to have rank zero as well. That is the reason why quadruply signed Selmer groups are the correct candidates which should relate to the quadruply signed p-adic L-functions (that interpolate linear combinations of critical L-values) we have defined above.

6.2. ***Quadruply signed main conjectures.*** We are now ready to state the quadruply signed main conjectures for the Rankin–Selberg convolutions of two p-nonordinary forms. We suppose throughout this section that the hypotheses of Proposition 5.3.4 are satisfied; in particular, we are assuming that Conjecture 5.3.1 holds.

Conjecture 6.2.1. *For $\mathfrak{S} = \{(\triangle, \square), (\bullet, \circ)\} \in \mathcal{S}$ and every character η of Γ_{tor}, the η-isotypic component $e_\eta \operatorname{Sel}_{\mathfrak{S}}(T^\vee(1)/\mathbb{Q}(\mu_{p^\infty}))$ of the quadruply signed Selmer group is $\mathcal{O}[\![\Gamma_1]\!]$-cotorsion and*

$$\operatorname{char}_{\mathcal{O}[\![\Gamma_1]\!]}\big(e_\eta \operatorname{Sel}_{\mathfrak{S}}(T^\vee(1)/\mathbb{Q}(\mu_{p^\infty}))^\vee\big) \;\big|\; (e_\eta \mathfrak{L}_{\mathfrak{S}})$$

as ideals of $\mathcal{O}[\![\Gamma_1]\!]$, with equality away from the support of $\operatorname{coker}(\mathrm{Col}_{\triangle,\square})$ *and* $\operatorname{coker}(\mathrm{Col}_{\bullet,\circ})$.

Remark 6.2.2. It is easy to prove that $e_\eta \mathfrak{L}_{\mathfrak{S}}$ is divisible by $\operatorname{char}\big(\operatorname{coker}(\mathrm{Col}_{\triangle,\square})\big)$ in $\mathbb{Q}_p \otimes \mathcal{O}[\![\Gamma_1]\!]$. This is the reason why the equality in the asserted divisibility in Conjecture 6.2.1 excludes the support of $\operatorname{coker}(\mathrm{Col}_{\triangle,\square})$. As illustrated by our main result (Corollary 7.4.9 below) towards the Pottharst-style analytic main conjectures, the error that is accounted by $\operatorname{coker}(\mathrm{Col}_{\triangle,\square})$ is inessential and it can be recovered.

On the other hand, the reason why we have to avoid the support of $\operatorname{coker}(\mathrm{Col}_{\bullet,\circ})$ is more subtle. In view of Proposition 5.3.2, we expect that the Euler system $\{\mathrm{BF}_{\bullet,\circ,m}\}$ be imprimitive when $\operatorname{coker}(e_\eta \mathrm{Col}_{\bullet,\circ})$ is not the unit ideal, in the sense that the bound it yields on the Selmer group will be off by $\operatorname{char}\big(\operatorname{coker}(e_\eta \mathrm{Col}_{\bullet,\circ})\big)$.

In this case, it is not clear to us whether or not it is possible to improve $\{\mathrm{BF}_{\bullet,\circ,m}\}$ to a primitive Euler system. These observations are also visible in Corollary 7.4.9 below.

In the remaining portions of this article we shall present evidence in favour of this conjecture. Until the end, we assume that $|k_f - k_g| \geq 3$ and $\eta = \omega^j$ for a fixed j such that $1 + (k_f + k_g)/2 < j \leq \max(k_f, k_g)$.

Proposition 6.2.3. *There exists a choice of $\mathfrak{S} \in \mathcal{S}$ such that $e_\eta \mathfrak{L}_\mathfrak{S} \neq 0$.*

Proof. Let $\mathcal{T} = \{\#, \flat\}^2$, and let $\mathscr{M}_{\mathrm{sign}}$ denote the 4×4 matrix, with rows and columns indexed by $\mathcal{T}$, whose (x, y) entry is $\mathrm{Col}_{x,\mathbb{Q}_p}(\mathrm{BF}_{y,1})$. Similarly, let $\mathcal{T}_{\mathrm{an}} = \{\alpha_f, \beta_f\} \times \{\alpha_g, \beta_g\}$, and let $\mathscr{M}_{\mathrm{an}}$ denote the 4×4 matrix whose x, y entry is $\mathcal{L}_x(\mathrm{BF}_{y,1})$. By Proposition 5.3.4, both matrices are antisymmetric (see Remark 6.1.3).

Among the six pairs of nondiagonal entries of $\mathscr{M}_{\mathrm{an}}$, four of them are given by p-adic Rankin–Selberg L-functions. By Corollary 3.8.1 and Remark 3.8.3, our hypotheses therefore imply that $e_\eta \mathscr{M}_{\mathrm{an}}$ is not the zero matrix.

However, our two matrices are related by the factorisation formula

$$\mathscr{M}_{\mathrm{an}} = (Q^{-1} M_{\log}) \cdot \mathscr{M}_{\mathrm{sign}} \cdot (Q^{-1} M_{\log})^{\,T},$$

so it follows that $e_\eta \mathscr{M}_{\mathrm{sign}}$ is also nonzero. Since the six pairs of nondiagonal entries of $\mathscr{M}_{\mathrm{sign}}$ are exactly the quadruply signed p-adic L-functions $\mathfrak{L}_\mathfrak{S}$, it follows that at least one of the $e_\eta \mathfrak{L}_\mathfrak{S}$ is nonzero as required. □

In order to apply the locally restricted Euler system argument devised in [Büyükboduk and Lei 2015, Appendix A], we will require the validity of the following hypothesis:

(**H.nA**) Neither $\bar{\rho}_f|_{G_{\mathbb{Q}_p}}$ nor $\bar{\rho}_f|_{G_{\mathbb{Q}_p}} \otimes \omega^{-1}$ is isomorphic to $\bar{\rho}_g^\vee|_{G_{\mathbb{Q}_p}}$, where ρ_f and ρ_g stand for Deligne's (cohomological) representations.

This assumption ensures that $H^0(\mathbb{Q}_p, \bar{T}) = H^2(\mathbb{Q}_p, \bar{T}) = 0$. We will also need to assume that

(**BI0**) $\epsilon_f \epsilon_g$ is nontrivial, $\gcd(N_f, N_g) = 1$.

as well as at least one of the following conditions:

(**BI1**) Neither f nor g is of CM type, and g has odd weight.

(**BI2**) f is not of CM-type, g is of CM-type and ϵ_g is neither the trivial character, nor the quadratic character attached to the CM field.

Thanks to [Loeffler 2017], when (**BI0**) and either (**BI1**) or (**BI2**) holds, one may choose a completion of the compositum of the Hecke fields of f and g (and set our coefficient ring $\mathcal{O}$ to be a finite flat extension of its ring of integers) in a way that the residue characteristic p of $\mathcal{O}$ is $> k_f + k_g + 2$ and the resulting Galois representation T verifies the following "Big Image" condition that is required to run the Euler system machinery:

- The residual representation $\bar{T}$ is absolutely irreducible.

- There exists an element $\tau \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}(\mu_{p^\infty}))$ such that $T/(\tau-1)T$ is a free $\mathcal{O}$-module of rank one.
- There exists an element $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}(\mu_{p^\infty}))$ which acts on T by multiplication by -1.

Theorem 6.2.4. *Suppose that $|k_f - k_g| \geq 3$ and $p > k_f + k_g + 2$. Assume the validity of Conjecture 5.3.1,* **(A–Sym)**, **(H.nA)**, **(BI0)** *and at least one of* **(BI1)**–**(BI2)**. *Suppose j is an integer with $1 + (k_f + k_g)/2 < j \leq \max(k_f, k_g)$ and choose $\mathfrak{S}$ that verifies the conclusion of Proposition 6.2.3 with $\eta = \omega^j$. Then the ω^j-isotypic component of the quadruply signed Selmer group $e_{\omega^j}\,\mathrm{Sel}_{\mathfrak{S}}(T^\vee(1)/\mathbb{Q}(\mu_{p^\infty}))$ is $\mathcal{O}[\![\Gamma_1]\!]$-cotorsion and*

$$e_{\omega^j}\mathfrak{L}_{\mathfrak{S}} \in \mathrm{char}_{\mathcal{O}[\![\Gamma_1]\!]}\big(e_{\omega^j}\,\mathrm{Sel}_{\mathfrak{S}}(T^\vee(1)/\mathbb{Q}(\mu_{p^\infty}))^\vee\big)$$

as ideals of $\mathcal{O}[\![\Gamma_1]\!] \otimes \mathbb{Q}_p$.

Remark 6.2.5. See Proposition 5.3.4 above where we provide a sufficient condition for the validity of **(A–Sym)**.

Proof of Theorem 6.2.4. This is a direct consequence of [Büyükboduk and Lei 2015, Theorem A.14], once we translate the language therein to our set up. Suppose $\mathfrak{S} = \{(\triangle, \square), (\bullet, \circ)\}$. Then the morphism Ψ in [loc. cit.] corresponds to the map

$$e_{\omega^j}\mathrm{Col}_{\triangle,\square} \oplus e_{\omega^j}\mathrm{Col}_{\bullet,\circ} : e_{\omega^j}H^1(\mathbb{Q}_p, \mathbb{T}) \to \mathcal{O}[\![\Gamma_1]\!]^{\oplus 2}.$$

The so-called Ψ-strict Selmer group that is denoted by $H^1_{\mathcal{F}_\Psi}(-,-)$ in [op. cit.] corresponds to our quadruply signed compact Selmer group $e_{\omega^j}H^1_{\mathcal{F}_{\mathfrak{S}}}(\mathbb{Q}, \mathbb{T})$ and its dual $H^1_{\mathcal{F}^*_\Psi}(-,-)$ to our quadruply signed discrete Selmer group $e_{\omega^j}\,\mathrm{Sel}_{\mathfrak{S}}(T^\vee(1)/\mathbb{Q}(\mu_{p^\infty}))$. Moreover, the integer g in [loc. cit.] equals to 2 in our case and the expression $\det([\Psi(\mathfrak{c}_i)]_{i=1}^g])$ in the statement of [Büyükboduk and Lei 2015, Theorem A.14(i)] is precisely $r_0^{-1}e_{\omega^j}\mathfrak{L}_{\mathfrak{S}}$ in our notation here (where $r_0 \in \mathbb{Z}$ is as in the formulation of Conjecture 5.3.1; note that since we have inverted p, this quantity does not make an appearance in the statement of our theorem).

Our running hypotheses guarantee the validity of all required assumptions for this result; only checking the validity of the condition (H.V) of [Büyükboduk and Lei 2015, Appendix A] (which translates in our setting to the condition that the quadruply signed compact Selmer group $e_{\omega^j}H^1_{\mathcal{F}_{\mathfrak{S}}}(\mathbb{Q}, \mathbb{T})$ be trivial) requires some work. The remainder of this proof is dedicated to show that the hypothesis (H.V) holds true in our set up.

We start with the reformulation of this condition. The Selmer group denoted by $H^1_{\mathcal{F}_L}(-,-)$ in [Büyükboduk and Lei 2015, Appendix A] corresponds to our Selmer group $e_{\omega^j}H^1_{\mathcal{F}_{\bullet,\circ}}(\mathbb{Q}, \mathbb{T})$ and the condition (H.V) requires that the map

$$e_{\omega^j}H^1_{\mathcal{F}_{\bullet,\circ}}(\mathbb{Q}, \mathbb{T}) \xrightarrow{(e_{\omega^j}\mathrm{Col}_{\triangle,\square} \oplus e_{\omega^j}\mathrm{Col}_{\bullet,\circ})\circ \mathrm{res}_p} \mathcal{O}[\![\Gamma_1]\!]^{\oplus 2}$$

be injective. By the defining property of $H^1_{\mathcal{F}_{\bullet,\circ}}(\mathbb{Q}, \mathbb{T})$, this is equivalent to checking that the map

$$e_{\omega^j}H^1_{\mathcal{F}_{\bullet,\circ}}(\mathbb{Q}, \mathbb{T}) \xrightarrow{e_{\omega^j}\mathrm{Col}_{\triangle,\square}\circ \mathrm{res}_p} \mathcal{O}[\![\Gamma_1]\!]$$

is injective. Since we already have $\mathrm{BF}_{\bullet,\circ,1} \in H^1_{\mathcal{F}_{\bullet,\circ}}(\mathbb{Q}, \mathbb{T})$ (as given by Conjecture 5.3.1, which we assume to hold) and we know (thanks to our choice of $\mathfrak{S}$) that

$$e_{\omega^j}\mathrm{Col}_{\triangle,\square} \circ \mathrm{res}_p(e_{\omega^j}\mathrm{BF}_{\bullet,\circ,1}) = e_{\omega^j}\mathfrak{L}_{\mathfrak{S}} \neq 0,$$

the condition (H.V) is equivalent to the requirement that $e_{\omega^j}H^1_{\mathcal{F}_{\bullet,\circ}}(\mathbb{Q}, \mathbb{T})$ has rank one. By Poitou–Tate global duality, this is in turn equivalent to checking that $e_{\omega^j}\,\mathrm{Sel}_{\bullet,\circ}(T^\vee(1)/\mathbb{Q}(\mu_{p^\infty}))$ is $\mathcal{O}[\![\Gamma_1]\!]$-cotorsion. We shall explain how to verify this fact.

Let us set $\mathbb{T}_1 := T \otimes \mathcal{O}[\![\Gamma_1]\!]^\iota$ (with diagonal Galois action). Choose a degree one polynomial $l \in \mathcal{O}[\![\Gamma_1]\!]$ that does not divide $e_{\omega^j}\mathfrak{L}_{\mathfrak{S}} \cdot \mathrm{char}\big(\mathrm{coker}(\mathrm{Col}_{\bullet,\circ})\big)$ and define $X := \mathbb{T}_1 \otimes \omega^{-j}/l$. Observe that we have (for $F = \mathbb{Q}_p$ or any finite abelian extension of $\mathbb{Q}$ unramified at p)

$$H^1(F, \mathbb{T}_1 \otimes \omega^{-j}) \xrightarrow{\sim} e_{\omega^j}H^1_{\mathrm{Iw}}(F(\mu_p), T) \tag{6.2.1}$$

by the inflation-restriction sequence. We define

$$H^1_{\mathcal{F}_{\bullet,\circ}}(\mathbb{Q}_p, \mathbb{T}_1 \otimes \omega^{-j}) \subset H^1(\mathbb{Q}_p, \mathbb{T}_1 \otimes \omega^{-j})$$

as the submodule that gets mapped isomorphically onto $e_{\omega^j}H^1_{\mathcal{F}_{\bullet,\circ}}(\mathbb{Q}, \mathbb{T})$ under the map (6.2.1). The isomorphism together with $e_{\omega^j}\mathrm{Col}_{\bullet,\circ}$ also induces a map

$$H^1(\mathbb{Q}_p, \mathbb{T}_1 \otimes \omega^{-j}) \to \mathcal{O}[\![\Gamma_1]\!]$$

(which we shall denote by the same symbol), whose kernel is precisely the submodule $H^1_{\mathcal{F}_{\bullet,\circ}}(\mathbb{Q}_p, \mathbb{T}_1 \otimes \omega^{-j})$.

For primes $\ell \neq p$, we shall also set $H^1_{\mathcal{F}_{\bullet,\circ}}(\mathbb{Q}_\ell, \mathbb{T}_1 \otimes \omega^{-j}) := H^1(\mathbb{Q}_\ell, \mathbb{T}_1 \otimes \omega^{-j})$ so that $\mathcal{F}_{\bullet,\circ}$ is a Selmer structure on $\mathbb{T}_1 \otimes \omega^{-1}$ in the sense of [Mazur and Rubin 2004]. It is easy to see that the dual Selmer group $H^1_{\mathcal{F}^*_{\bullet,\circ}}(\mathbb{Q}, \mathbb{T}_1^\vee(1) \otimes \omega^j)^\vee$ is $\mathcal{O}[\![\Gamma_1]\!]$-torsion if and only if $e_{\omega^j}\,\mathrm{Sel}_{\bullet,\circ}(T^\vee(1)/\mathbb{Q}(\mu_{p^\infty}))^\vee$ is $\mathcal{O}[\![\Gamma_1]\!]$-torsion. Thanks to [Mazur and Rubin 2004, Lemma 3.5.3], our claim that $e_{\omega^j}\,\mathrm{Sel}_{\bullet,\circ}(T^\vee(1)/\mathbb{Q}(\mu_{p^\infty}))$ is cotorsion follows once we verify that

$$H^1_{\mathcal{F}^*_{\bullet,\circ}}(\mathbb{Q}, \mathbb{T}_1^\vee(1) \otimes \omega^j)[l] \cong H^1_{\mathcal{F}^*_{\bullet,\circ}}(\mathbb{Q}, X^\vee(1))$$

has finite cardinality. Note that we have written $\mathcal{F}_{\bullet,\circ}$ (resp. $\mathcal{F}^*_{\bullet,\circ}$) for the propagation of the Selmer structure $\mathcal{F}_{\bullet,\circ}$ to X (resp. for the Selmer structure on $X^\vee(1)$ dual to $\mathcal{F}_{\bullet,\circ}$ on X).

Consider the following diagram with exact rows and Cartesian squares:

$$\begin{array}{ccccccc}
0 & \longrightarrow & H^1_{\mathcal{F}_{\bullet,\circ}}(\mathbb{Q}_p, \mathbb{T}_1 \otimes \omega^{-j}) & \longrightarrow & H^1(\mathbb{Q}_p, \mathbb{T}_1 \otimes \omega^{-j}) & \xrightarrow{\mathrm{Col}_{\bullet,\circ}} & \mathcal{O}[\![\Gamma_1]\!] \\
 & & \downarrow & & \downarrow & & \downarrow \\
0 & \longrightarrow & H^1_{\mathcal{F}}(\mathbb{Q}_p, X) & \longrightarrow & H^1(\mathbb{Q}_p, X) & \xrightarrow{\phi} & \mathcal{O}
\end{array}$$

where the vertical arrows are induced by reduction modulo l (which we henceforth denote by π_X); the map ϕ is defined by the Cartesian square on the right; the submodule $H^1_{\mathcal{F}}(\mathbb{Q}_p, X)$ by the exactness of the

second row and the dotted arrow by chasing the diagram. Note that the map ϕ is not the zero map thanks to our choice of l.

For $\ell \neq p$, let us also define

$$H^1_{\mathcal{F}}(\mathbb{Q}_\ell, X) := \ker\big(H^1(\mathbb{Q}_\ell, X) \to H^1(\mathbb{Q}_\ell^{\mathrm{ur}}, X \otimes \mathbb{Q}_p)\big).$$

It follows from [Mazur and Rubin 2004, Lemma 5.3.1(i)] (together with the dotted arrow in the diagram above) that the reduction map modulo l induces a map

$$H^1_{\mathcal{F}_{\bullet,\circ}}(\mathbb{Q}_\ell, \mathbb{T}_1 \otimes \omega^{-j}) \to H^1_{\mathcal{F}}(\mathbb{Q}_\ell, X) \tag{6.2.2}$$

for every prime ℓ (including p), which factors through the injection $H^1_{\mathcal{F}_{\bullet,\circ}}(\mathbb{Q}_\ell, X) \subset H^1_{\mathcal{F}}(\mathbb{Q}_\ell, X)$. In particular, we have $\mathcal{F}_{\bullet,\circ} \leq \mathcal{F}$ in the sense of [Mazur and Rubin 2004, Definition 2.1.1] and we have an injective map

$$\mathrm{KS}(X, \mathcal{F}_{\bullet,\circ}) \hookrightarrow \mathrm{KS}(X, \mathcal{F}) \tag{6.2.3}$$

between the corresponding modules of Kolyvagin systems.

Using Lemma 3.7.1 of [op. cit.], it follows that the Selmer structure $\mathcal{F}$ is cartesian (in the sense of [Mazur and Rubin 2004, Definition 1.1.4]). Moreover, it is easy to see (recalling that the map ϕ is not the zero map) that the core Selmer rank of $\mathcal{F}$ equals one. In particular, by [Mazur and Rubin 2004, Corollary 5.2.13], the finiteness of $H^1_{\mathcal{F}}(\mathbb{Q}_\ell, X)$ is equivalent to exhibiting a single Kolyvagin system in $\mathrm{KS}(X, \mathcal{F})$, whose initial term is nonzero. We shall prove that the Euler system $\{e_{\omega^j}\mathrm{BF}_{\bullet,\circ,n}\}$ of doubly signed Beilinson–Flach classes descend to a Kolyvagin system with this property and this completes the proof.

Let us write $\mathrm{BF}^{\omega^j}_{\bullet,\circ} \in H^1(\mathbb{Q}(\mu_n), \mathbb{T}_1 \otimes \omega^{-j})$ for the class that maps to $e_{\omega^j}\mathrm{BF}_{\bullet,\circ,n}$ under the isomorphism (6.2.1) and let us set $\mathrm{BF}^{\omega^j}_{\bullet,\circ} := \{\mathrm{BF}^{\omega^j}_{\bullet,\circ,n}\}$. As we have explained in Remark 5.3.3, $\{\mathrm{BF}^{\omega^j}_{\bullet,\circ}\}$ is a locally restricted Euler system. Moreover, [Büyükboduk and Lei 2015, Theorem A.11] applies (with our choices here, recall that the Selmer structure $\mathcal{F}_L$ in [loc. cit.] corresponds to our $\mathcal{F}_{\bullet,\circ}$) and produces a Kolyvagin system

$$\kappa^{\bullet,\circ} = \{\kappa_n^{\bullet,\circ}\} \in \mathrm{KS}(\mathbb{T}_1 \otimes \omega^{-j}, \mathcal{F}_{\bullet,\circ}).$$

We remark that the Kolyvagin system $\kappa^{\bullet,\circ}$ is obtained from the Euler system $\mathrm{BF}^{\omega^j}_{\bullet,\circ,n}$ via [Mazur and Rubin 2004, Theorem 5.3.3]. However, the results of Mazur and Rubin show a priori only that $\kappa^{\bullet,\circ} \in \mathrm{KS}(\mathbb{T}_1 \otimes \omega^{-j}, \mathcal{F}_\Lambda)$ (namely, $\kappa^{\bullet,\circ}$ is a Kolyvagin system for the canonical Selmer structure $\mathcal{F}_\Lambda$ of [Mazur and Rubin 2004, Definition 5.3.2]). The fact that the classes $\kappa_n^{\bullet,\circ}$ verify the required local conditions at p follows from the fact that $\mathrm{BF}^{\omega^j}_{\bullet,\circ}$ is in fact locally restricted, as explained in detail in the proof of [Büyükboduk and Lei 2015, Theorem A.11].

On projecting $\kappa^{\bullet,\circ}$ via π_X and composing with the injection (6.2.3), we obtain a Kolyvagin system $\pi_X(\kappa^{\bullet,\circ}) \in \mathrm{KS}(X, \mathcal{F})$. We have

$$\pi_X(\kappa^{\bullet,\circ})_1 = \pi_X(\kappa_1^{\bullet,\circ}) = \pi_X(\mathrm{BF}^{\omega^j}_{\bullet,\circ,1})$$

for the initial term of this Kolyvagin system. We are reduced to prove that $\pi_X(\mathrm{BF}^{\omega^j}_{\bullet,\circ,1}) \neq 0$ for the choices above.

We now recall that $l \nmid e_{\omega^j}\mathfrak{L}_{\mathfrak{S}} = \mathrm{Col}_{\triangle,\square} \circ \mathrm{res}_p(\mathrm{BF}^{\omega^j}_{\bullet,\circ,1})$ by the choice we made on the degree one polynomial l. This in turn implies that $\pi_X(\mathrm{BF}^{\omega^j}_{\bullet,\circ,1}) \neq 0$, as required. □

Remark 6.2.6. In the proof of Theorem 6.2.4, we only needed the antisymmetry property in **(A–Sym)** to hold for $\{(\triangle, \square), (\bullet, \circ)\}$ with $m = 1$ and $\{(\bullet, \circ), (\bullet, \circ)\}$ for all $m \in \mathcal{N}(\mathcal{P})$.

7. Analytic main conjectures

Our goal in this section is to translate our results on signed Iwasawa main conjectures into the "analytic" language of Pottharst and Benois (see [Pottharst 2013; Benois 2015]). This gives main conjectures which directly involve the p-adic Rankin–Selberg L-functions (3.6.1), which is advantageous since the interpolating properties of these L-functions are much more explicit than those of the signed p-adic L-functions $\mathfrak{L}_{\mathfrak{S}}$ of the previous section. However, it has the disadvantage of throwing away all p-torsion information.

7.1. *Cohomology of* (φ, Γ)*-modules.*

For each $0 \leq r < 1$, let

$$\mathrm{ann}(r, 1) := \{x \in \mathbb{C}_p : r \leq |x|_p < 1\}.$$

For E the finite extension of $\mathbb{Q}_p$ in Section 1, we define the *Robba ring*

$$\mathcal{R}_E := \left\{ f(\pi) = \sum_{n=-\infty}^{\infty} a_n \pi^n \in E[\![\pi]\!] \;\middle|\; f(\pi) \text{ converges on } \mathrm{ann}(r, 1) \text{ for some } r \right\}.$$

The Robba ring comes equipped with actions of Γ and the Frobenius φ, via the same formulae as in Section 2 above.

Definition 7.1.1. A (φ, Γ)-module over $\mathcal{R}_E$ is a free module of finite rank d endowed with a semilinear Frobenius φ such that $\mathrm{Mat}(\varphi) \in \mathrm{GL}_d(\mathcal{R}_E)$ and with a continuous commuting semilinear action of Γ.

There exists a functor $V \mapsto \mathbb{D}^{\dagger}_{\mathrm{rig}}(V)$ between the category of p-adic representations of $G_p := \mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$ with coefficients in E and the category of (φ, Γ)-modules over $\mathcal{R}_E$ (see [Cherbonnier and Colmez 1999; Fontaine 1990; Berger 2002]).

For any (φ, Γ)-module $\mathbb{D}$, we define its *analytic Iwasawa cohomology* $H^{\bullet}_{\mathrm{an}}(\mathbb{D})$ to be the cohomology of the complex

$$\left[\mathbb{D} \xrightarrow{\psi - 1} \mathbb{D}\right],$$

where ψ is the left inverse of φ and the terms are placed in degrees 1 and 2 respectively. If V is a p-adic representation of G_p, then the results of [Pottharst 2013] show that

$$H^i_{\mathrm{an}}(\mathbb{Q}_p(\mu_{p^\infty}), V) \cong H^i_{\mathrm{an}}(\mathbb{D}^{\dagger}_{\mathrm{rig}}(V)),$$

where $H^i_{\mathrm{an}}(\mathbb{Q}_p(\mu_{p^\infty}), V) = H^i(\mathbb{Q}_p, \mathbb{V}^\dagger)$ denotes the analytic Iwasawa cohomology of V as defined in Section 2.4.

7.2. ***Selmer complexes.*** Now let V be a continuous E-linear representation of $G_{\mathbb{Q}}$ which is unramified at almost all primes. Following Benois and Pottharst, we now recall the complexes defined in Section 2.3 of [Benois 2015], which are an "analytic" version of the Selmer complexes of Nekovář [2006]. Letting $\mathbb{D}$ be any (φ, Γ)-submodule of $\mathbb{D}^\dagger_{\mathrm{rig}}(V|_{G_{\mathbb{Q}_p}})$, we obtain a *Selmer complex* $S(\mathbb{Q}, \mathbb{V}^\dagger; \mathbb{D})$ in the derived category of $\mathcal{H}$-modules. This is defined by a mapping fibre

$$\mathrm{cone}\left[C^\bullet(G_{\mathbb{Q},\Sigma}, \mathbb{V}^\dagger) \oplus \bigoplus_{v\in\Sigma} U_v^+ \xrightarrow{\mathrm{res}_v - i_v^+} \bigoplus_{v\in\Sigma} C^\bullet(G_{\mathbb{Q}_v}, \mathbb{V}^\dagger)\right][-1],$$

where Σ is any sufficiently large finite set of places, and the U_v^+ are appropriate "local condition" complexes; we choose these to be the unramified local conditions for $v \neq p$, and the local condition at p is given by the analytic Iwasawa cohomology of the submodule $\mathbb{D} \subseteq \mathbb{D}^\dagger_{\mathrm{rig}}(V|_{G_{\mathbb{Q}_p}})$. We write $H^\bullet(\mathbb{Q}, \mathbb{V}^\dagger; \mathbb{D})$ for the cohomology groups of the Selmer complex.

We shall apply this with $V = R_f^* \otimes R_g^*$ and with local conditions $\mathbb{D}$ given by φ-stable subspaces of $\mathbb{D}_{\mathrm{cris}}(V)$. We set

$$\begin{aligned}\mathbb{D}_{\mathrm{cris}}(V)^\lambda &:= \mathbb{D}_{\mathrm{cris}}(V)^{\lambda_f\lambda_g} \oplus \mathbb{D}_{\mathrm{cris}}(V)^{\lambda_f\mu_g} \\ &= \mathbb{D}_{\mathrm{cris}}(R_f^*)^{\lambda_f} \otimes \mathbb{D}_{\mathrm{cris}}(R_g^*)\end{aligned}$$

and

$$\mathbb{D}_{\mathrm{cris}}(V)^{\lambda,\mu} := \mathbb{D}_{\mathrm{cris}}(R_f^*)^{\lambda_f} \otimes \mathbb{D}_{\mathrm{cris}}(R_g^*) + \mathbb{D}_{\mathrm{cris}}(R_f^*) \otimes \mathbb{D}_{\mathrm{cris}}(R_g^*)^{\mu_g}.$$

Let $\mathbb{D}_{\lambda,\mu}$ and $\mathbb{D}_\lambda$ be the (φ, Γ)-submodules of $\mathbb{D}^\dagger_{\mathrm{rig}}(T)$ corresponding to $\mathbb{D}_{\mathrm{cris}}(T)^\lambda$ and $\mathbb{D}_{\mathrm{cris}}(T)^{\lambda,\mu}$ respectively (see [Berger 2008] for an explicit description). We note that $\mathbb{D}_{\lambda,\mu}$ and $\mathbb{D}_\lambda$ play the role of *regular submodules* defined by Benois [2015].

7.3. ***Analytic main conjectures.*** We are now ready to state the *analytic Iwasawa main conjecture* formulated by Benois and Pottharst, specialised to our setting.

Conjecture 7.3.1. *The module $H^2(\mathbb{Q}, \mathbb{V}^\dagger; \mathbb{D}_\lambda)$ is torsion, and its characteristic ideal is given by*

$$\mathrm{char}_{\mathcal{H}} H^2(\mathbb{Q}, \mathbb{V}^\dagger; \mathbb{D}_\lambda) = L_p(f_\lambda, g) \cdot \mathcal{H},$$

where $L_p(f_\lambda, g)$ denotes the geometric p-adic L-function attached to f and g.

(One can similarly define a Selmer group and formulate a main conjecture for the "extra" p-adic L-functions $L_p^?$ introduced above, but we shall not give the details here.) We now proceed to show how our results in the previous sections on signed Selmer groups aid us to deduce some partial results towards Conjecture 7.3.1.

7.4. ***Bounds for analytic Selmer groups.*** Throughout the remainder of this section, the hypotheses of Theorem 6.2.4 are in effect. We begin by reformulating the result of Theorem 6.2.4 as a bound for $H^2(\mathbb{Q}, \mathbb{T})$; we shall then translate this into a bound for the analytic Selmer group.

Proposition 7.4.1. *There exists a choice* $(\bullet, \circ)$ *and* $(\triangle, \square)$ *such that neither* $e_\eta \mathrm{Col}_{\triangle,\square} \circ \mathrm{res}_p(\mathrm{BF}_{\bullet,\circ,1})$ *nor* $e_\eta \mathrm{Col}_{\bullet,\circ} \circ \mathrm{res}_p(\mathrm{BF}_{\triangle,\square,1})$ *is zero.*

Proof. This follows from Propositions 6.2.3 and 5.3.4. □

From now on, we fix $(\bullet, \circ)$ and $(\triangle, \square)$ such that $e_\eta \mathrm{Col}_{\triangle,\square} \circ \mathrm{res}_p(\mathrm{BF}_{\bullet,\circ,1}) \neq 0$.

As in Proposition 3.9.2, this implies that $e_\eta H^1(\mathbb{Q}, \mathbb{T})$ is free of rank two. Choose an $e_\eta\Lambda$-basis $\{c_1, c_2\}$ of this module.

Definition 7.4.2. Let $r_1, r_2, D \in \Lambda$ be nonzero elements such that

$$E \cdot \mathrm{BF}_{\bullet,\circ,1} = r_1\big(\mathrm{Col}_{\bullet,\circ} \circ \mathrm{res}_p(c_1)c_2 - \mathrm{Col}_{\bullet,\circ} \circ \mathrm{res}_p(c_2)c_1\big)$$
$$E \cdot \mathrm{BF}_{\triangle,\square,1} = r_2\big(\mathrm{Col}_{\triangle,\square} \circ \mathrm{res}_p(c_1)c_2 - \mathrm{Col}_{\triangle,\square} \circ \mathrm{res}_p(c_2)c_1\big).$$

Here, the first equality takes place in $H^1_{\mathcal{F}_{\bullet,\circ}}(\mathbb{Q}, \mathbb{T})$, whereas the second in $H^1_{\mathcal{F}_{\triangle,\square}}(\mathbb{Q}, \mathbb{T})$.

Note that E, r_1, r_2 with the required properties exist since both modules $H^1_{\mathcal{F}_{\bullet,\circ}}(\mathbb{Q}, \mathbb{T})$ and $H^1_{\mathcal{F}_{\triangle,\square}}(\mathbb{Q}, \mathbb{T})$ have rank one.

Lemma 7.4.3.
$$r_1 = -r_2.$$

Proof. This is an immediate consequence of Proposition 5.3.4. □

Proposition 7.4.4. $E \cdot \mathrm{char}\big(H^2(\mathbb{Q}, \mathbb{T})\big)$ *divides* $r_1 \cdot \mathrm{char}\big(\mathrm{coker}(\mathrm{Col}_{\bullet,\circ})\big)$.

Proof. Let us set $H^1_{/\bullet,\circ}(\mathbb{Q}_p, \mathbb{T}) := H^1(\mathbb{Q}_p, \mathbb{T})/H^1_{\mathcal{F}_{\bullet,\circ}}(\mathbb{Q}_p, \mathbb{T})$ and write $\mathrm{res}^s_{\bullet,\circ}$ for the compositum of the arrows

$$H^1(\mathbb{Q}, \mathbb{T}) \xrightarrow{\mathrm{res}_p} H^1(\mathbb{Q}_p, \mathbb{T}) \to H^1_{/\bullet,\circ}(\mathbb{Q}_p, \mathbb{T}).$$

Poitou–Tate global duality gives rise to the following five-term exact sequence:

$$0 \to H^1_{\mathcal{F}_{\bullet,\circ}}(\mathbb{Q}, \mathbb{T})/\Lambda \cdot \mathrm{BF}_{\bullet,\circ,1} \to H^1(\mathbb{Q}, \mathbb{T})/(\mathrm{BF}_{\bullet,\circ,1}, \mathrm{BF}_{\triangle,\square,1}) \to \frac{H^1_{/\bullet,\circ}(\mathbb{Q}_p, \mathbb{T})}{\mathrm{res}^s_{\bullet,\circ}(\mathrm{BF}_{\triangle,\square,1})}$$
$$\to H^1_{\mathcal{F}^*_{\bullet,\circ}}(\mathbb{Q}, \mathbb{T}^\vee(1))^\vee \to H^2(\mathbb{Q}, \mathbb{T}) \to 0. \quad (7.4.1)$$

The locally restricted Euler system machinery shows that

$$\mathrm{char}\big(H^1_{\mathcal{F}^*_{\bullet,\circ}}(\mathbb{Q}, \mathbb{T}^\vee(1))^\vee\big) \mid \mathrm{char}\big(H^1_{\mathcal{F}_{\bullet,\circ}}(\mathbb{Q}, \mathbb{T})/\Lambda \cdot \mathrm{BF}_{\bullet,\circ,1}\big). \quad (7.4.2)$$

Combining (7.4.1) with the divisibility (7.4.2), we infer that

$$\mathrm{char}\big(H^2(\mathbb{Q}, \mathbb{T})\big)\mathrm{char}\big(\mathrm{coker}(\mathrm{Col}_{\bullet,\circ})\big)^{-1}\mathrm{Col}_{\bullet,\circ} \circ \mathrm{res}_p(\mathrm{BF}_{\triangle,\square,1})$$
$$\text{divides } \mathrm{char}\left(\frac{H^1(\mathbb{Q}, \mathbb{T})}{(\mathrm{BF}_{\bullet,\circ,1}, \mathrm{BF}_{\triangle,\square,1})}\right). \quad (7.4.3)$$

Moreover, we have

$$\mathrm{Col}_{\bullet,\circ}(\mathrm{BF}_{\triangle,\square,1}) = E^{-1} r_2 \det \circ \mathrm{Col}(\triangle, \square, \bullet, \circ) \tag{7.4.4}$$

where we have set

$$\det \circ \mathrm{Col}(\triangle, \square, \bullet, \circ) := \det \begin{pmatrix} \mathrm{Col}_{\triangle,\square} \circ \mathrm{res}_p(c_1) & \mathrm{Col}_{\bullet,\circ} \circ \mathrm{res}_p(c_1) \\ \mathrm{Col}_{\triangle,\square} \circ \mathrm{res}_p(c_2) & \mathrm{Col}_{\bullet,\circ} \circ \mathrm{res}_p(c_2) \end{pmatrix}.$$

Let $f_{\triangle,\square} := \mathrm{Col}_{\triangle,\square} \circ \mathrm{res}_p$ (and we similarly define $f_{\bullet,\circ}$). Note that since we have

$$\mathrm{char}\Big(E(c_1, c_2) \Big/ \big(r_1(f_{\bullet,\circ}(c_1)c_2 - f_{\bullet,\circ}(c_2)c_1),\ r_2(f_{\triangle,\square}(c_1)c_2 - f_{\triangle,\square}(c_2)c_1) \big) \Big) = E^{-2} r_1 r_2 \det \circ \mathrm{Col}(\triangle, \square, \bullet, \circ)\,. \tag{7.4.5}$$

it follows from definitions that

$$\mathrm{char}\left(\frac{H^1(\mathbb{Q}, \mathbb{T})}{(\mathrm{BF}_{\bullet,\circ,1}, \mathrm{BF}_{\triangle,\square,1})} \right) = E^{-2} r_1 r_2 \det \circ \mathrm{Col}(\triangle, \square, \bullet, \circ). \tag{7.4.6}$$

Combining (7.4.3), (7.4.4) and (7.4.6), the asserted divisibility follows. □

Let us choose $F, s_1, s_2 \in \mathcal{H} \setminus \{0\}$ so that

$$F \cdot \mathrm{BF}_{\lambda,\mu} = s_1 \big(\mathcal{L}_{\lambda,\mu} \circ \mathrm{res}_p(c_1)c_2 - \mathcal{L}_{\lambda,\mu} \circ \mathrm{res}_p(c_2)c_1 \big)$$
$$F \cdot \mathrm{BF}_{\lambda,\mu'} = s_2 \big(\mathcal{L}_{\lambda,\mu'} \circ \mathrm{res}_p(c_1)c_2 - \mathcal{L}_{\lambda,\mu'} \circ \mathrm{res}_p(c_2)c_1 \big).$$

We also set

$$\det \circ \mathcal{L}(\lambda, \mu, \lambda, \mu') := \det \begin{pmatrix} \mathcal{L}_{\lambda,\mu} \circ \mathrm{res}_p(c_1) & \mathcal{L}_{\lambda,\mu'} \circ \mathrm{res}_p(c_1) \\ \mathcal{L}_{\lambda,\mu} \circ \mathrm{res}_p(c_2) & \mathcal{L}_{\lambda,\mu'} \circ \mathrm{res}_p(c_2) \end{pmatrix}.$$

Lemma 7.4.5. $s_1 = -s_2$.

Proof. This follows from Theorem 3.9.1. □

Proposition 7.4.6. *We have the following divisibility of $\mathcal{H}$-ideals*:

$$\mathrm{char}\big(H^2(\mathbb{Q}, \mathbb{V}^\dagger; \mathbb{D}_{\lambda,\mu}) \big) \mid \frac{r_1 F}{s_1 E} \cdot \mathrm{char}\big(\mathrm{coker}\, \mathrm{Col}_{\bullet,\circ} \big) \cdot \mathrm{char}\left(\frac{H^1(\mathbb{Q}, \mathbb{V}^\dagger; \mathbb{D}_{\lambda,\mu})}{\mathcal{H} \cdot \mathrm{BF}_{\lambda,\mu,1}} \right).$$

Proof. The proof of this assertion is essentially the proof of Proposition 7.4.4 in reverse, starting off with the 5-term exact sequence of $\mathcal{H}$-modules

$$0 \to \frac{H^1(\mathbb{Q}, \mathbb{V}^\dagger; \mathbb{D}_{\lambda,\mu})}{\mathcal{H} \cdot \mathrm{BF}_{\lambda,\mu,1}} \to \frac{H^1(\mathbb{Q}, \mathbb{V}^\dagger)}{\mathcal{H} \cdot \mathrm{BF}_{\lambda,\mu,1} + \mathcal{H} \cdot \mathrm{BF}_{\lambda,\mu',1}} \to \frac{H^1_{/\lambda,\mu}(\mathbb{Q}_p, \mathbb{V}^\dagger)}{\mathrm{res}^s_p(\mathrm{BF}_{\lambda,\mu'})} \to H^2(\mathbb{Q}, \mathbb{V}^\dagger; \mathbb{D}_{\lambda,\mu}) \to H^2(\mathbb{Q}, \mathbb{V}^\dagger) \to 0,$$

where $H^1_{/\lambda,\mu}(\mathbb{Q}_p, \mathbb{V}^\dagger) := H^1(\mathbb{Q}_p, \mathbb{V}^\dagger)/H^1_{\mathrm{an}}(\mathbb{Q}_p, \mathbb{D}_{\lambda,\mu})$ and res^s_p is the natural map

$$\mathrm{res}^s_p : H^1(\mathbb{Q}, \mathbb{V}) \to H^1_{/\lambda,\mu}(\mathbb{Q}_p, \mathbb{V}^\dagger).$$

Notice that we also rely on the fact that $\mathcal{L}_{\lambda,\mu} : H^1_{/\lambda,\mu}(\mathbb{Q}_p, \mathbb{V}^\dagger) \to \mathcal{H}$ is surjective, and the fact that $H^2(\mathbb{Q}, \mathbb{V}^\dagger) = \mathcal{H} \otimes_\Lambda H^2(\mathbb{Q}, \mathbb{T})$ when we use here the statement of Proposition 7.4.4. □

Corollary 7.4.7. *The ideal* $\mathrm{char}\big(H^2(\mathbb{Q}, \mathbb{V}^\dagger; \mathbb{D}_\lambda)\big)$ *divides the ideal*

$$\mathrm{char}\big(\mathrm{coker}\,\mathrm{Col}_{\bullet,\circ}\big)E^{-1}r_1 F s_1^{-1}\mathcal{L}_{\lambda,\mu'}(\mathrm{BF}_{\lambda,\mu,1}).$$

Proposition 7.4.8.

$$r_1 F = r_0 s_1 E.$$

Proof. Recall the element $\mathrm{BF}_1 \in \mathrm{Frac}\,\mathcal{H} \otimes_\Lambda \bigwedge^2 H^1_{\mathrm{Iw}}(\mathbb{Q}(\mu_{p^\infty}), T)$ given as in Theorem 3.9.1. Let us choose $h \in \mathrm{Frac}(\mathcal{H})^\times$ so that we have $e_\eta \mathrm{BF}_1 = h \cdot (c_1 \wedge c_2)$. It follows from (5.3.1) together with the defining property of E and r_1 that $r_0 h = r_1/E$. Comparing the definition of h and s_1/F, we also see that $h = s_1/F$. Our assertion follows. □

Corollary 7.4.9. *The ideal* $\mathrm{char}\big(H^2(\mathbb{Q}, \mathbb{V}^\dagger; \mathbb{D}_\lambda)\big)$ *divides*

$$\mathrm{char}\big(\mathrm{coker}\,\mathrm{Col}_{\bullet,\circ}\big)L_p(f_\lambda, g).$$

Proof. This follows on combining Corollary 7.4.7 and Proposition 7.4.8, together with the observation that $r_0 \in \mathbb{Z}$ is invertible in the ring $\mathcal{H}$. □

Appendix: Images of Coleman maps

We describe the images of various Coleman maps up to pseudo-isomorphisms, which is relevant to our discussion in Remark 6.2.2. Throughout, T denotes the representation $R_f^* \otimes R_g^*$ and

$$\mathcal{L}_T : H^1_{\mathrm{Iw}}(\mathbb{Q}_p, T) \to \mathcal{H} \otimes \mathbb{D}_{\mathrm{cris}}(T)$$

as before. We recall the following result from [Lei et al. 2011].

Proposition A.1. *Let* $z \in H^1_{\mathrm{Iw}}(\mathbb{Q}_p, T)$, δ *a Dirichlet character of conductor* $p^n > 1$ *and* $0 \le j \le k_f + k_g + 1$, *then*

$$(1 - p^j\varphi)^{-1}(1 - p^{-j-1}\varphi^{-1})\chi^j(\mathcal{L}_T(z)) \in \mathrm{Fil}^{-j}\,\mathbb{D}_{\mathrm{cris}}(T),$$
$$(p^j\varphi)^{-n}\big(\chi^j\delta(\mathcal{L}_T(z))\big) \in \mathbb{Q}_p(\mu_{p^{n-1}}) \otimes \mathrm{Fil}^{-j}\,\mathbb{D}_{\mathrm{cris}}(T).$$

Proof. This is [Lei et al. 2011, Proposition 4.8]. □

Given a character $\eta \in \hat{\Delta}$ and an integer $0 \le j \le k_f + k_g + 1$, we define

$$V_{\eta,j} := \begin{cases} (1 - p^j\varphi)(\varphi - p^{-j-1})^{-1}\,\mathrm{Fil}^{-j}\,\mathbb{D}_{\mathrm{cris}}(T) & \text{if } \eta = \omega^j, \\ \mathrm{Fil}^{-j}\,\mathbb{D}_{\mathrm{cris}}(T) & \text{otherwise.} \end{cases}$$

Via our chosen basis $\{v_1, v_2, v_3, v_4\}$ of $\mathbb{D}_{\mathrm{cris}}(T)$, we identify $\mathbb{D}_{\mathrm{cris}}(T) \otimes \mathbb{Q}_p$ with L. Let us write $\underline{\mathrm{Col}} : H^1_{\mathrm{Iw}}(\mathbb{Q}_p, T) \to \mathcal{O} \otimes \Lambda^{\oplus 4}$ for the morphism given by $(\mathrm{Col}_{\#,\#}, \mathrm{Col}_{\#,\flat}, \mathrm{Col}_{\flat,\#}, \mathrm{Col}_{\flat,\flat})$. For each $\eta \in \hat{\Delta}$, we may then identify $e_\eta\underline{\mathrm{Col}}$ as a map landing inside $\mathcal{O}[\![X]\!]^{\oplus 4}$ as before via $1 + X = \gamma$, where γ is our chosen topological generator of Γ_1. Recall that $u = \chi(\gamma)$.

Corollary A.2. *Let* $z \in H^1_{\mathrm{Iw}}(\mathbb{Q}_p, T)$, $0 \le j \le k_f + k_g + 1$ *and* $\eta \in \hat{\Delta}$. *Then*,

$$e_\eta\underline{\mathrm{Col}}(z)|_{X = u^j - 1} \in V_{\eta,j}.$$

Proof. Note that $e_\eta \mathcal{L}_T(z)|_{X=u^j-1} = \chi^j(\eta\omega^{-j})(\mathcal{L}_T(z))$. Therefore, Proposition A.1 says that

$$e_\eta \mathcal{L}_T(z)|_{X=u^j-1} \in P_{\eta,j}(\mathrm{Fil}^{-j}\,\mathbb{D}_{\mathrm{cris}}(T)),$$

where $P_{\eta,j}$ is given by

$$\begin{cases} (1-p^j\varphi)(1-p^{-j-1}\varphi^{-1})^{-1} & \text{if } \eta = \omega^j, \\ vp & \text{otherwise.} \end{cases}$$

Recall from (5.1.1) that

$$\mathcal{L}_T(z) = \begin{pmatrix} v_1 & v_2 & v_3 & v_4 \end{pmatrix} \cdot M_{\log} \cdot \begin{pmatrix} \mathrm{Col}_{\#,\#}(z) \\ \mathrm{Col}_{\#,\flat}(z) \\ \mathrm{Col}_{\flat,\#}(z) \\ \mathrm{Col}_{\flat,\flat}(z) \end{pmatrix}.$$

Note that $e_\eta M_{\log}|_{X=u^j-1} = A$, which is a consequence of (4.2.3). Recall that A is the matrix of φ with respect to the basis $\{v_1, v_2, v_3, v_4\}$, which implies our result. □

Following [Lei et al. 2011, Proposition 4.11], this allows us to deduce the following description of the image of $\underline{\mathrm{Col}}$.

Corollary A.3. *Let $\eta \in \hat{\Delta}$, then*

$$e_\eta \mathrm{Im}(\underline{\mathrm{Col}}) \otimes E = \left\{ \underline{F} \in \mathcal{O}[\![X]\!] \otimes E : \underline{F}(u^j - 1) \in V_{\eta,j},\ 0 \le j \le k_f + k_g + 1 \right\}.$$

If $\mathfrak{S} = \{(\triangle, \square), (\bullet, \circ)\} \in \mathcal{S}$, this corresponds to a two-dimensional subspace in $\mathbb{D}_{\mathrm{cris}}(T) \otimes E$, generated by two elements of the basis $\{v_1, v_2, v_3, v_4\}$, which we denote by $V_{\mathfrak{S}}$. If we write $\mathrm{Col}_{\mathfrak{S}}$ for the wedge product

$$\mathrm{Col}_{\triangle,\square} \wedge \mathrm{Col}_{\bullet,\circ} : \bigwedge^2 H^1_{\mathrm{Iw}}(\mathbb{Q}_p, T) \to \Lambda,$$

then we may describe its image by

$$e_\eta \mathrm{Im}(\mathrm{Col}_{\mathfrak{S}}) \otimes E = \prod_{j=0}^{k_f+k_g+1} (X - u^j + 1)^{n_{\mathfrak{S},\eta,j}} \mathcal{O}[\![X]\!] \otimes E,$$

where $n_{\mathfrak{S},\eta,j} = \dim_E V_{\mathfrak{S}} \cap V_{\eta,j}$. If we do not tensor our image by E, we have that $e_\eta \mathrm{Im}(\mathrm{Col}_{\mathfrak{S}})$ is pseudo-isomorphic to $p^{\mu_{\mathfrak{S},\eta}} \prod_{j=0}^{k_f+k_g+1} (X - u^j + 1)^{n_{\mathfrak{S},\eta,j}} \mathcal{O}[\![X]\!]$ for some integer $\mu_{\mathfrak{S},\eta}$.

Acknowledgement

The authors would like to thank the referee for reading an earlier version of the manuscript and for valuable suggestions, which led to many improvements of the paper.

References

[Benois 2015] D. Benois, "Selmer complexes and p-adic Hodge theory", pp. 36–88 in *Arithmetic and geometry*, edited by L. Dieulefait et al., London Math. Soc. Lecture Note Ser. **420**, Cambridge Univ. Press, 2015. MR Zbl

[Berger 2002] L. Berger, "Représentations p-adiques et équations différentielles", *Invent. Math.* **148**:2 (2002), 219–284. MR Zbl

[Berger 2003] L. Berger, "Bloch and Kato's exponential map: three explicit formulas", *Doc. Math.* Extra Volume: Kazuya Kato's fiftieth birthday (2003), 99–129. MR Zbl

[Berger 2004] L. Berger, "Limites de représentations cristallines", *Compos. Math.* **140**:6 (2004), 1473–1498. MR Zbl

[Berger 2008] L. Berger, "Équations différentielles p-adiques et (ϕ, N)-modules filtrés", pp. 13–38 in *Représentations p-adiques de groupes p-adiques, I: Représentations galoisiennes et* (ϕ, Γ)*-modules*, Astérisque **319**, Société Mathématique de France, Paris, 2008. MR Zbl

[Büyükboduk and Lei 2015] K. Büyükboduk and A. Lei, "Coleman-adapted Rubin–Stark Kolyvagin systems and supersingular Iwasawa theory of CM abelian varieties", *Proc. Lond. Math. Soc.* (3) **111**:6 (2015), 1338–1378. MR Zbl

[Büyükboduk and Lei 2016] K. Büyükboduk and A. Lei, "Iwasawa theory of elliptic modular forms over imaginary quadratic fields at non-ordinary primes", preprint, 2016. arXiv

[Büyükboduk et al. 2018] K. Büyükboduk, A. Lei, and G. Venkat, "Iwasawa theory for Symmetric Square of non-p-ordinary eigenforms", preprint, 2018. arXiv

[Cherbonnier and Colmez 1999] F. Cherbonnier and P. Colmez, "Théorie d'Iwasawa des représentations p-adiques d'un corps local", *J. Amer. Math. Soc.* **12**:1 (1999), 241–268. MR Zbl

[Fontaine 1990] J.-M. Fontaine, "Représentations p-adiques des corps locaux, I", pp. 249–309 in *The Grothendieck Festschrift Volume 2*, edited by P. Cartier et al., Progr. Math. **87**, Birkhäuser, Boston, 1990. MR Zbl

[Kings et al. 2015] G. Kings, D. Loeffler, and S. L. Zerbes, "Rankin–Eisenstein classes for modular forms", preprint, 2015. To appear in *Amer. J. Math.* arXiv

[Kings et al. 2017] G. Kings, D. Loeffler, and S. L. Zerbes, "Rankin–Eisenstein classes and explicit reciprocity laws", *Camb. J. Math.* **5**:1 (2017), 1–122. MR Zbl

[Kobayashi 2003] S. Kobayashi, "Iwasawa theory for elliptic curves at supersingular primes", *Invent. Math.* **152**:1 (2003), 1–36. MR Zbl

[Lei 2011] A. Lei, "Iwasawa theory for modular forms at supersingular primes", *Compos. Math.* **147**:3 (2011), 803–838. MR Zbl

[Lei 2017] A. Lei, "Bounds on the Tamagawa numbers of a crystalline representation over towers of cyclotomic extensions", *Tohoku Math. J.* (2) **69**:4 (2017), 497–524. MR Zbl

[Lei et al. 2010] A. Lei, D. Loeffler, and S. L. Zerbes, "Wach modules and Iwasawa theory for modular forms", *Asian J. Math.* **14**:4 (2010), 475–528. MR Zbl

[Lei et al. 2011] A. Lei, D. Loeffler, and S. L. Zerbes, "Coleman maps and the p-adic regulator", *Algebra Number Theory* **5**:8 (2011), 1095–1131. MR Zbl

[Lei et al. 2014] A. Lei, D. Loeffler, and S. L. Zerbes, "Euler systems for Rankin–Selberg convolutions of modular forms", *Ann. of Math.* (2) **180**:2 (2014), 653–771. MR Zbl

[Lei et al. 2017] A. Lei, D. Loeffler, and S. L. Zerbes, "On the asymptotic growth of Bloch–Kato–Shafarevich–Tate groups of modular forms over cyclotomic extensions", *Canad. J. Math.* **69**:4 (2017), 826–850. MR Zbl

[Loeffler 2014] D. Loeffler, "p-adic integration on ray class groups and non-ordinary p-adic L-functions", pp. 357–378 in *Iwasawa theory 2012*, edited by T. Bouganis and O. Venjakob, Contrib. Math. Comput. Sci. **7**, Springer, 2014. MR Zbl

[Loeffler 2017] D. Loeffler, "Images of adelic Galois representations for modular forms", *Glasg. Math. J.* **59**:1 (2017), 11–25. MR Zbl

[Loeffler and Zerbes 2016a] D. Loeffler and S. L. Zerbes, "Iwasawa theory for the symmetric square of a modular form", *J. Reine Angew. Math.* (online publication December 2016).

[Loeffler and Zerbes 2016b] D. Loeffler and S. L. Zerbes, "Rankin–Eisenstein classes in Coleman families", *Res. Math. Sci.* **3**:26 (2016). Zbl

[Mazur and Rubin 2004] B. Mazur and K. Rubin, *Kolyvagin systems*, Mem. Amer. Math. Soc. **799**, 2004. MR Zbl

[Nekovář 2006] J. Nekovář, *Selmer complexes*, Astérisque **310**, Société Mathématique de France, Paris, 2006. MR Zbl

[Otsuki 2009] R. Otsuki, "Construction of a homomorphism concerning Euler systems for an elliptic curve", *Tokyo J. Math.* **32**:1 (2009), 253–278. MR Zbl

[Perrin-Riou 1994] B. Perrin-Riou, "Théorie d'Iwasawa des représentations p-adiques sur un corps local", *Invent. Math.* **115**:1 (1994), 81–161. MR Zbl

[Perrin-Riou 1998] B. Perrin-Riou, "Systèmes d'Euler p-adiques et théorie d'Iwasawa", *Ann. Inst. Fourier* (*Grenoble*) **48**:5 (1998), 1231–1307. MR Zbl

[Pottharst 2013] J. Pottharst, "Analytic families of finite-slope Selmer groups", *Algebra Number Theory* **7**:7 (2013), 1571–1612. MR Zbl

[Rubin 1996] K. Rubin, "A Stark conjecture "over $\mathbf{Z}$" for abelian L-functions with multiple zeros", *Ann. Inst. Fourier* (*Grenoble*) **46**:1 (1996), 33–62. MR Zbl

[Urban 2017] E. Urban, "Application to the three variable Rankin–Selberg p-adic L-functions", 2017. Appendix II in F. Andreatta and A Iovita, "Triple product p-adic L-functions associated to finite slope p-adic families of modular forms", preprint, 2017. arXiv

Communicated by Christopher Skinner
Received 2018-02-12 Revised 2018-09-13 Accepted 2019-02-10

kazim.buyukboduk@ucd.ie *UCD School of Mathematics and Statistics, University College Dublin, Ireland*

antonio.lei@mat.ulaval.ca *Département de Mathématiques et de Statistique, Université Laval, Pavillion Alexandre-Vachon, Quebec City, QC, Canada*

d.a.loeffler@warwick.ac.uk *Mathematics Institute, University of Warwick, Zeeman Building, Coventry, United Kingdom*

guhanvenkat.harikumar.1@ulaval.ca *Département de Mathématiques et de Statistique, Laval University, Pavillion Alexandre-Vachon, Quebec City, QC, Canada*

msp

ALGEBRA AND NUMBER THEORY 13:4 (2019)
dx.doi.org/10.2140/ant.2019.13.943

Cycle integrals of modular functions, Markov geodesics and a conjecture of Kaneko

Paloma Bengoechea and Özlem Imamoglu

In this paper we study the values of modular functions at the Markov quadratics which are defined in terms of their cycle integrals along the associated closed geodesics. These numbers are shown to satisfy two properties that were conjectured by Kaneko. More precisely we show that the values of a modular function f, along any branch B of the Markov tree, converge to the value of f at the Markov number which is the predecessor of the tip of B. We also prove an interlacing property for these values.

1. Introduction

A well known theorem of Dirichlet asserts that for any irrational number x, there are infinitely many rational numbers p/q satisfying $|x - p/q| < 1/q^2$. For irrational numbers that are algebraic, thanks to a theorem of Roth [1955], the exponent 2 is optimal. The constant factor, on the other hand, can be improved and a classical theorem of Hurwitz asserts that for every irrational number x there exist infinitely many rational numbers p/q satisfying

$$\left|x - \frac{p}{q}\right| < \frac{1}{\sqrt{5}q^2}.$$

The constant $1/\sqrt{5}$ is best possible but if we exclude as x the numbers that are $\mathrm{PGL}(2,\mathbb{Z})$-equivalent to the golden ratio $(1+\sqrt{5})/2$, the constant $1/\sqrt{5}$ improves to $1/\sqrt{8}$. If we also exclude the numbers that are $\mathrm{PGL}(2,\mathbb{Z})$-equivalent to $\sqrt{2}$, then the constant improves to $5/\sqrt{221}$. By proceeding in this way, one obtains the Lagrange spectrum defined by

$$L := \{\nu(x)\}_{x\in\mathbb{R}} \subseteq [0, 1/\sqrt{5}] \quad \text{with} \quad \nu(x) = \liminf_{q\to\infty} q\|qx\|,$$

where $\|x\|$ denotes the distance from a real number x to a closest integer. The quantity $\nu(x)$ provides a measure of approximation of x by the rationals. For almost all $x \in \mathbb{R}$ we have $\nu(x) = 0$ and when $\nu(x) > 0$ we call x badly approximable. Real quadratic irrationals are badly approximable, the worst ones being the golden ratio and its $\mathrm{PGL}(2,\mathbb{Z})$-equivalents, followed by $\sqrt{2}$ and its $\mathrm{PGL}(2,\mathbb{Z})$-equivalents, etc. The Lagrange spectrum is not discrete (see [Hall 1947]) but the part of the spectrum in the subinterval $\left(\frac{1}{3}, 1/\sqrt{5}\right]$ corresponding to classes of worst irrational numbers is, with $\frac{1}{3}$ as its only accumulation point.

Bengoechea's research is supported by SNF grant 173976.

MSC2010: primary 11F03; secondary 11J06.

Keywords: modular forms, cycle integrals, markov numbers, j-invariant.

$L\cap\left(\frac{1}{3}, 1/\sqrt{5}\right]$ is well understood thanks to the work of Markov [1879; 1880] which connects this question of Diophantine approximation to the Diophantine equation

$$x^2+y^2+z^2=3xyz. \tag{1}$$

The set of Markov triples comprising the positive integer solutions (x, y, z) of (1) can be obtained starting with $(1, 1, 1)$, $(1, 1, 2)$, $(2, 1, 5)$ and then proceeding recursively going from (x, y, z) to the new triples obtained by Vieta involutions $(z, y, 3yz-x)$ and $(x, z, 3xz-y)$. The Markov numbers are the greatest coordinates of Markov triples. They form the Markov sequence

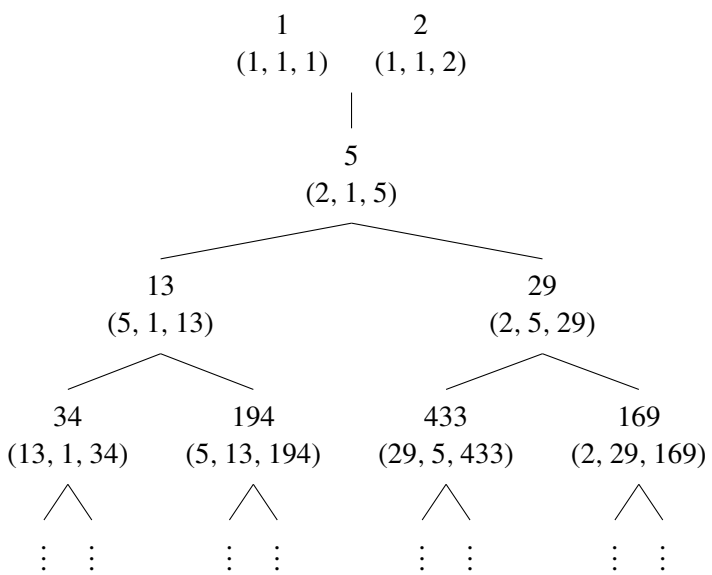
$$\{m_i\}_{i=1}^{\infty}=\{1, 2, 5, 13, 29, 34, 89, 169, 194, \ldots\}.$$

The Markov number m_i is associated to a quadratic irrationality

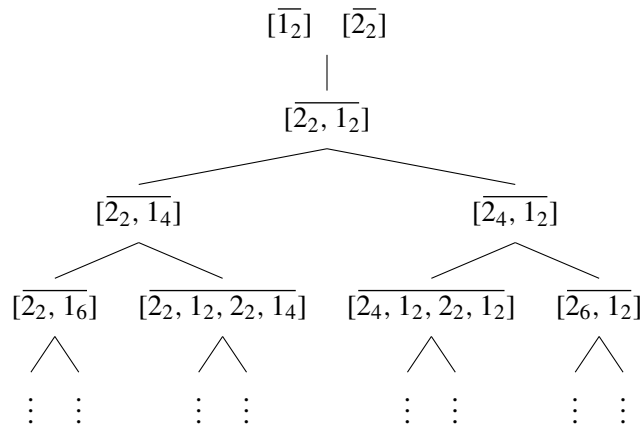
$$\theta_i=\frac{3m_i-2k_i+\sqrt{9m_i^2-4}}{2m_i},$$

where k_i is an integer that satisfies $a_ik_i\equiv b_i \pmod{m_i}$ and (a_i, b_i, m_i) is a solution to (1) with m_i maximal. Since k_i is uniquely defined modulo m_i, θ_i is uniquely defined modulo 1. Markov showed that $\nu(\theta_i)=\sqrt{9-4/m_i^2}$, and $L\cap\left(\frac{1}{3}, 1/\sqrt{5}\right]=\{\nu(\theta_i)\}_{i\geq 1}$. Moreover, any $x\in\mathbb{R}$ for which $\nu(x)\in L\cap\left(\frac{1}{3}, 1/\sqrt{5}\right]$ is $\mathrm{PGL}(2, \mathbb{Z})$-equivalent to a Markov quadratic θ_i.

Markov numbers come with a tree structure, inherited from Vieta involutions, that arranges them as



Here (a, b, c) is a solution to (1). The Markov quadratics inherit the same tree structure which can be given in terms of their continued fractions as



where b_n means that b is repeated n times. We note that it is more convenient to write $[\overline{1_2}]$ instead of $[\bar{1}]$ in connection with the conjunction operator in (5). The fact that all of the partial quotients of Markov quadratics are 1 or 2 and many of their other properties can be found in [Aigner 2013; Bombieri 2007; Malyshev 1977] (See for example Corollary 1.27 in [Aigner 2013].)

Markov numbers arise in many different contexts: see [Bourgain et al. 2016b; 2016a; Ghosh and Sarnak 2017] for some recent developments regarding the Markov surfaces.

The main goal of this paper is to study the values of modular functions along the tree associated to the Markov quadratics.

Let $\Gamma = \mathrm{PSL}(2, \mathbb{Z})$. For a general quadratic irrationality $w \in \mathbb{Q}(\sqrt{D})$ and a modular function f for Γ, the "value" of f at w is defined in terms of the integral of f along the geodesic cycle $C_w \subset \Gamma\backslash\mathcal{H}$ associated to w. More precisely

$$f(w) := \int_{C_w} f(z)\, ds,$$

where ds is the hyperbolic arc length. We can normalize the number $f(w)$ by the length of the geodesic C_w and define

$$f^{\mathrm{nor}}(w) := \frac{f(w)}{2 \log \varepsilon_D},$$

where ε_D is the fundamental unit (see Section 2A).

The values of modular functions at real quadratic irrationalities were introduced in [Duke et al. 2011] and independently in [Kaneko 2009]. In [Duke et al. 2011] their averages over ideal classes were shown to be coefficients of mock modular forms whereas Kaneko [2009] studied their individual values $f^{\mathrm{nor}}(w)$ (in the case that the modular function is the Klein's j invariant), and based on numerical calculations he made several interesting observations and conjectures.

In this paper we prove two of Kaneko's conjectures which involve the values of modular functions at the Markov quadratics. Let B be any branch of the Markov tree where with a branch we mean a path on the tree without any zigzags. Our first theorem shows that if w_n^B is the n-th Markov quadratic on a branch B and w_0^B is the predecessor of the tip of B then the normalized values $f^{\mathrm{nor}}(w_n^B)$, for any modular

function f, converge to the value $f^{\mathrm{nor}}(w_0^B)$. (For more precise definitions of the tip of a branch and its predecessor see Section 3A.) More precisely:

Theorem 1.1. *Let f be a modular function defined on $\mathcal{H}$. For any branch B of the Markov tree we have*

$$\lim_{n\to\infty} f^{\mathrm{nor}}(w_n^B) = f^{\mathrm{nor}}(w_0^B).$$

Our second theorem proves an eventual monotonicity result which also partially proves the interlacing property of the values for the Markov quadratics that was conjectured by Kaneko.

Theorem 1.2. *Let f be a modular function on $\mathcal{H}$, let B be any branch of the Markov tree. Then there exists a constant $N_{f,B}$ such that, for all $n \geq N_{f,B}$, the real and imaginary parts of $f^{\mathrm{nor}}(w_{n+1}^B)$ lie between the real and respectively imaginary parts of $f^{\mathrm{nor}}(w_0^B)$ and $f^{\mathrm{nor}}(w_n^B)$.*

The rest of the paper is organized as follows. In the next section we give the preliminaries about cycle integrals and continued fractions. In Section 3, we give the basic properties of the Markov quadratics and the Markov tree. In Sections 4 and 5 we study the values of modular functions on the Markov tree and prove Theorems 1.1 and 1.2 respectively.

2. Preliminaries

2A. ***Cycle integrals.*** Let w be a real quadratic irrationality and $\tilde{w}$ be its conjugate, so w and $\tilde{w}$ are the roots of a quadratic equation

$$ax^2 + bx + c = 0 \qquad (a, b, c \in \mathbb{Z}, \quad (a, b, c) = 1)$$

with discriminant $D = b^2 - 4ac > 0$. We change $[a, b, c]$ to $-[a, b, c]$ if necessary and write

$$w = \frac{-b + \sqrt{D}}{2a}, \quad \tilde{w} = \frac{-b - \sqrt{D}}{2a}.$$

The geodesic S_w in $\mathcal{H}$ joining w and $\tilde{w}$ is given by the equation

$$a|z|^2 + b\,\mathrm{Re}(z) + c = 0 \qquad (z \in \mathcal{H}).$$

The stabilizer Γ_w of w in Γ preserves the quadratic form $Q_w = [a, b, c]$, and hence S_w. The group Γ_w is infinite cyclic; it corresponds to the group U_D^2 of units of norm one of $\mathbb{Q}(\sqrt{D})$ via the isomorphism:

$$\Gamma_w \longrightarrow U_D^2, \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto (a - cw)^2. \tag{2}$$

We denote by A_w the generator of Γ_w,

$$A_w = \begin{pmatrix} \frac{1}{2}(t - bu) & -cu \\ au & \frac{1}{2}(t + bu) \end{pmatrix},$$

where (t, u) is the smallest positive solution to Pell's equation $t^2 - Du^2 = 4$, and we denote by ε the generator of the infinite cyclic part of U_D whose square corresponds to A_w by the isomorphism (2).

For any modular function f, since the group Γ_w preserves the expression $f(z)Q_w(z,1)^{-1}\,dz$, one can define the cycle integral of f along $C_w = S_w/\Gamma_w$, also viewed as the "value" of f at w, by the complex number

$$f(w) := \int_{C_w} \frac{\sqrt{D}f(z)}{Q_w(z,1)}\,dz. \tag{3}$$

The factor $\sqrt{D}$ is introduced here for convenience but is also natural since with the constant function $f \equiv 1$, (3) gives the length of the geodesic C_w. The integral defining $f(w)$ is Γ-invariant and can in fact be taken along any path in $\mathcal{H}$ from z_0 to $A_w^{-1}z_0$, where z_0 is any point in $\mathcal{H}$. Note that this gives an orientation on S_w from w to $\tilde{w}$, which is counterclockwise if $a > 0$ and clockwise if $a < 0$. We normalize the number $f(w)$ by the length of the geodesic C_w which is given by

$$\int_{C_w} \frac{\sqrt{D}}{Q_w(z,1)}\,dz = 2\log\varepsilon$$

and we define the normalized value as

$$f^{\mathrm{nor}}(w) := \frac{f(w)}{2\log\varepsilon}.$$

2B. ***The "+" and "−" continued fractions.*** Let $(b_0, b_1, b_2, \ldots)$ denote the "−" continued fraction

$$(b_0, b_1, b_2, \ldots) = b_0 - \cfrac{1}{b_1 - \cfrac{1}{b_2 - \cfrac{1}{\ddots}}}$$

and $[a_0, a_1, a_2, \ldots]$ be the "+" continued fraction

$$[a_0, a_1, a_2, \ldots] = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots}}}.$$

Every real number w has a "−" continued fraction expansion $w = (b_0, b_1, b_2, \ldots)$ with $b_i \in \mathbb{Z}$ and $b_i \geq 2$ for $i \geq 1$ and a unique "+" continued fraction expansion $w = [a_0, a_1, a_2, \ldots]$ with $a_i \in \mathbb{Z}$ and $a_i \geq 1$ for $i \geq 1$. The "−" continued fraction expansion of w is obtained by setting $w_0 = w$ and inductively $b_i = \lceil w_i \rceil$, $w_{i+1} = 1/(b_i - w_i) = ST^{-b_i}(w_i)$, where $S(x) = -1/x$ and $T(x) = x + 1$. The "+" continued fraction expansion is obtained by setting $a_i = \lfloor w_i \rfloor$, $w_{i+1} = 1/(w_i - a_i) = \varepsilon T^{-a_i}(w_i)$, where $\varepsilon(x) = 1/x$. Hence the "−" continued fraction is given by transformations of Γ on the real line, whereas the "+" continued fraction corresponds to transformations of $\mathrm{GL}(2,\mathbb{Z})$. To go from the "+" to the "−" continued fraction expansions, the general rule is

$$[a_0, a_1, a_2, \ldots] \to \big(a_0 + 1, \underbrace{2, \ldots, 2}_{a_1 - 1}, a_2 + 2, \underbrace{2, \ldots, 2}_{a_3 - 1}, a_4 + 2, \ldots\big). \tag{4}$$

It is well known that a real number w is a quadratic irrationality if and only if its "$-$" continued fraction expansion (or equivalently, its "$+$" continued fraction) is eventually periodic:

$$w = (b_0, b_1, \ldots, b_k, \overline{b_{k+1}, \ldots, b_{k+r}}),$$

where the line over $b_{k+1}, \ldots, b_{k+r}$ denotes the period. We say that w is *purely periodic* when all the partial quotients repeat. It will be useful for the rest of the paper to remember the following statements:

(I) Two quadratic irrationalities have the same "$-$" period if and only if they are Γ-equivalent.

(II) w has a purely periodic "$-$" continued fraction expansion if and only if $0 < \tilde{w} < 1 < w$, where $\tilde{w}$ is the conjugate of w.

(III) If $w = (\overline{b_0, \ldots, b_r})$, then $1/\tilde{w} = (\overline{b_r, \ldots, b_0})$.

These statements and more information about negative continued fractions can be found in [Zagier 1981, p. 126 ff].

The following lemma gives an upper bound for the distance between two real numbers in terms of the number of first partial quotients for which they coincide.

Lemma 2.1. *If the "$-$" continued fraction expansions of u and v coincide in the first $r+1$ partial quotients and their "$+$" continued fraction expansions have only 1's and 2's, then*

$$|u - v| \le 10\left(\frac{2}{1+\sqrt{5}}\right)^{2r}.$$

Proof. Let u and v be as in the statement of the lemma. Then one can see, by applying the rule (4), that also the "$+$" continued fraction expansions of u and v coincide in the first $r+1$ partial quotients. Hence, if we set $a_0, \ldots, a_r$ to be those partial quotients, the rational number $p/q = [a_0, \ldots, a_r]$ is a convergent of both u and v. Then it is well known that

$$\left|u - \frac{p}{q}\right| \le \frac{1}{q^2}, \quad \left|v - \frac{p}{q}\right| \le \frac{1}{q^2}$$

and

$$q \ge \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^{r}.$$

Therefore,

$$|u - v| \le \left|u - \frac{p}{q}\right| + \left|v - \frac{p}{q}\right| \le 10\left(\frac{2}{1+\sqrt{5}}\right)^{2r}. \qquad \square$$

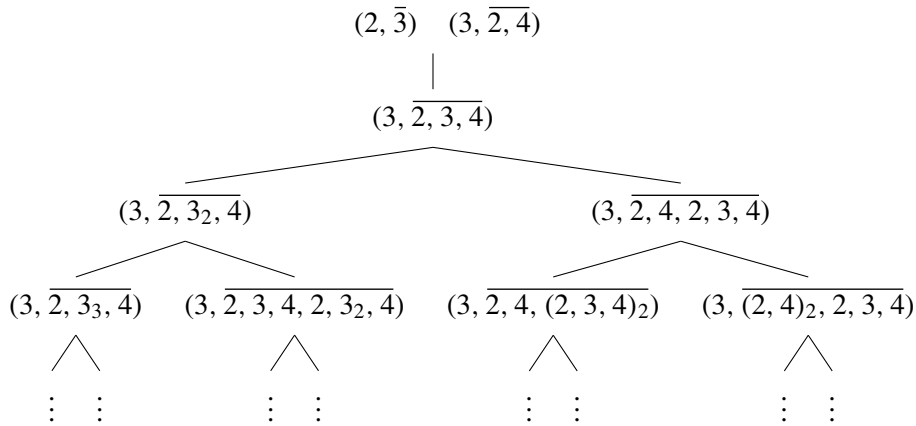
3. Markov Tree

3A. ***Markov's quadratics.*** Let $\{m_i\}_{i=1}^{\infty} = \{1, 2, 5, 13, 29, 34, 89, 169, 194, \ldots\}$ be the set of Markov numbers. As in the introduction, for each Markov number m_i, we let

$$\theta_i = \frac{3m_i - 2k_i + \sqrt{9m_i^2 - 4}}{2m_i}$$

be the Markov quadratic where k_i is an integer that satisfies $a_i k_i \equiv b_i \pmod{m_i}$ and (a_i, b_i, m_i) is a solution to (1) with m_i maximal. Changing the representative for $k_i \mod m_i$ does not change the Γ orbit of θ_i. In Markov's theory, only $\mathrm{PGL}(2,\mathbb{Z})$-equivalence classes are relevant, which implies that the order of (a_i, b_i) does not matter. Since we need Γ-equivalence, which distinguishes nonreal $f(\theta_i)$ and its conjugate, here the order of (a_i, b_i) becomes relevant. We fix it so that $\mathrm{Im}(f(w)) > 0$.

The Markov tree $\mathcal{T}$ associated to the Markov quadratics given in the introduction is in terms of the "+" continued fractions. Since the cycle integrals are Γ and not $\mathrm{PGL}(2,\mathbb{Z})$ invariant, we will rather work with the "−" continued fraction. By following the rule (4), the Markov tree $\mathcal{T}$ becomes in the "−" continued fraction



Note that each branch (a path with no zigzags) in the tree $\mathcal{T}$ comes with a left or right orientation. We call a branch a left (right) branch if starting from its first vertex on the top and going downwards the branch leans towards left (right). Since no zigzag paths are allowed, each branch has a unique orientation. For example, the branch with the quadratics $(3, \overline{2,3,4})$, $(3, \overline{2,3_2,4})$, $(3, \overline{2,3_3,4})$ is a left branch, whereas the branch with $(3, \overline{2,3,4})$, $(3, \overline{2,4,2,3,4})$, $(3, \overline{(2,4)_2,2,3,4})$ is a right branch. We call the first vertex at the top of any branch its tip. Except for the two singular cases of $(2, \overline{3})$ and $(3, \overline{2,4})$, each Markov number lies both on a right and a left branch but it is the tip of only a left or a right branch, except for $(3, \overline{2,3,4})$ which is the tip of both the leftmost and the rightmost branches.

In the case of "+" continued fractions we consider a *conjunction operation* of two periods as

$$[\overline{s_0, \ldots, s_n}] \odot [\overline{t_0, \ldots, t_m}] = [\overline{s_0, \ldots, s_n, t_0, \ldots, t_m}]. \tag{5}$$

All Markov quadratics can be constructed by using this operation, starting with $[\overline{1}_2]$ and $[\overline{2}_2]$. Indeed, each Markov quadratic is the result of the conjunction operation of its predecessor on the same branch and the predecessor of the tip of the branch.

For the "−" continued fraction, the rule is also the conjunction of periods except for the leftmost branch, where the n-th Markov quadratic is $(3, \overline{2,3_n,4})$. Indeed, let $x = [\overline{s_0, \ldots, s_n}] = (b_0, \overline{b_1, \ldots, b_k})$ and $y = [\overline{t_0, \ldots, t_m}] = (c_0, \overline{c_1, \ldots, c_\ell})$. For any branch different from the rightmost branch, by applying (4) together with the observation that $s_n = t_m = 1$ are in odd positions, so they do not contribute in the

"−" expansion, we obtain

$$x \odot y = (b_0, \overline{b_1, \ldots, b_{k-1}, t_0 + 2, c_1, \ldots, c_{\ell-1}, s_0 + 2}).$$

But t_0 is equal to 1 on the leftmost branch and 2 on any other branch, and $s_0 = 2$. For the rightmost branch, (4) also gives

$$x \odot y = (b_0, \overline{b_1, \ldots, b_{k-1}, 4, c_1, \ldots, c_{\ell-1}, s_0 + 2})$$

and $s_0 = 2$.

Throughout the paper, we denote by w_n^B $(n \geq 1)$ the n-th Markov quadratic on a branch B of the tree and w_0^B the left (right) predecessor of the tip w_1^B of B if B is a left (right) branch. For example, if $B = L$ is the leftmost branch, then $w_0^L = (2, \overline{3})$, $w_1^L = (3, \overline{2, 3, 4})$, $w_2^L = (3, \overline{2, 3_2, 4})$, $w_3^L = (3, \overline{2, 3_3, 4})$, etc. If $B = R$ is the rightmost branch, then $w_0^R = (3, \overline{2, 4})$, $w_1^R = (3, \overline{2, 3, 4})$, $w_2^R = (3, \overline{2, 4, 2, 3, 4})$, $w_3^R = (3, \overline{(2, 4)_2, 2, 3, 4})$, etc.

The n-th Markov quadratic on a left branch $B \neq L$ can be written as

$$w_n^B = (3, \overline{a_1, \ldots, a_s, (b_1, \ldots, b_r)_n}), \tag{6}$$

where $w_0^B = (3, \overline{b_1, \ldots, b_r})$ and $a_1, \ldots, a_s$ depend only on B. On a right branch B, we have

$$w_n^B = (3, \overline{(b_1, \ldots, b_r)_{n-1}, a_1, \ldots, a_s}), \tag{7}$$

and on the leftmost branch L we have

$$w_n^L = (3, \overline{2, 3_n, 4}). \tag{8}$$

Remark 3.1. The leftmost branch in the Markov tree is also called the Fibonacci branch since the associated Markov numbers on this branch are the odd indexed Fibonacci numbers. Similarly the rightmost branch is associated with the odd indexed Pell numbers which are defined by the recurrence $P_0 = 0$, $P_1 = 1$ and $P_{n+1} = 2P_n + P_{n-1}$ (see [Aigner 2013, p. 49]).

3B. ***The cycle of quadratics of a Markov number.*** For any quadratic irrationality w, it is known that the hyperbolic element A_w is conjugate to a word in T and V, where

$$T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad V = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

If in particular $w = w_n^B$ is a quadratic on $\mathcal{T}$ $(n \geq 0)$, then the associated hyperbolic element $A_{w_n^B}$ can be written as a word in T and V. More specifically, $A_{w_n^B} = A_0^{-1} \cdots A_{\ell_n}^{-1}$, where $A_0 = I$ and $A_i \in \{T^{-1}, V^{-1}\}$ for $1 \leq i \leq \ell_n$ are given by the algorithm:

$$w_{n,0}^B = w_n^B, \qquad w_{n,i+1}^B = A_{i+1}(w_{n,i}^B) \quad (i \geq 0),$$

where

$$A_{i+1} = \begin{cases} T^{-1} & \text{if } \lfloor w_{n,i}^B \rfloor \geq 1, \\ V^{-1} & \text{otherwise.} \end{cases}$$

Hence

$$w_{n,i}^B = A_i \cdots A_0 w_n^B, \quad i = 0, \ldots, \ell_n, \tag{9}$$

and ℓ_n is the length of the word $A_{w_n^B}$, or equivalently, the length of the cycle of quadratics $\{w_{n,i}^B\}_i$ of w_n^B. As the following example demonstrates, this procedure applied to a Markov quadratic in fact cycles back and hence terminates.

Example 3.2. For example, the cycle of $w_1^L = (3, \overline{2, 3, 4})$ on the leftmost branch is:

$$\begin{aligned}
w_{1,0}^L &= (3, \overline{2, 3, 4}),\\
w_{1,1}^L &= T^{-1}(w_{1,0}^L) = (2, \overline{2, 3, 4}),\\
w_{1,2}^L &= T^{-1}(w_{1,1}^L) = (1, \overline{2, 3, 4}),\\
w_{1,3}^L &= V^{-1}(w_{1,2}^L) = (1, \overline{3, 4, 2}),\\
w_{1,4}^L &= V^{-1}(w_{1,3}^L) = (2, \overline{4, 2, 3}),\\
w_{1,5}^L &= T^{-1}(w_{1,4}^L) = (1, \overline{4, 2, 3}),\\
w_{1,6}^L &= V^{-1}(w_{1,5}^L) = (3, \overline{2, 3, 4}) = w_{1,0}^L.
\end{aligned}$$

The length is $\ell_1 = 6$ and $A_{w_1^L} = ITTVVTV$.

From now on we restrict to a left branch but not the leftmost branch. All the following arguments apply in the same way if B is a right branch or $B = L$, the leftmost branch. The small difference in the arguments arise due to the different conjunction operations necessary, which are given in (7) for the right and in (8) for the leftmost branches.

We now consider w_n^B, in a left branch $B \neq L$, written as in (6). Then

$$\ell_n = n\ell_0 + \sum_{i=1}^{s}(a_i - 1), \tag{10}$$

where

$$\ell_0 = \sum_{i=1}^{r}(b_i - 1)$$

is the length of the cycle of w_0^B. The number of partial quotients in the period of w_1^B is $s + r$ and the conjunction operation ensures that this is $\leq 2r$. Hence $s \leq r$ and since $a_i \leq 4$, we have

$$\ell_n \leq 3r(n+1). \tag{11}$$

It is convenient to set

$$a = \sum_{i=1}^{s}(a_i - 1)$$

and

$$\boldsymbol{p} = (b_1, \ldots, b_r), \quad \boldsymbol{p}_k = (b_1, \ldots, b_r)_k, \quad \boldsymbol{q}_k = (b_r, \ldots, b_1)_k,$$

where the subindex k means that the continued fraction is repeated k times. With these notations, the cycle of w_n^B is of the form:

$$\begin{aligned}
w_{n,0}^B &= (3, \overline{a_1, \dots, a_s, \boldsymbol{p}_n}),\\
w_{n,1}^B &= (2, \overline{a_1, \dots, a_s, \boldsymbol{p}_n}),\\
w_{n,2}^B &= (1, \overline{a_1, \dots, a_s, \boldsymbol{p}_n}),\\
w_{n,3}^B &= (a_1 - 1, \overline{a_2, \dots, a_s, \boldsymbol{p}_n, a_1}),\\
&\vdots\\
w_{n,a}^B &= (3, \overline{\boldsymbol{p}_n, a_1, \dots, a_s}),\\
&\vdots\\
w_{n,a+\ell_0}^B &= (3, \overline{\boldsymbol{p}_{n-1}, a_1, \dots, a_s, \boldsymbol{p}}),\\
&\vdots\\
w_{n,a+n\ell_0}^B &= (3, \overline{a_1, \dots, a_s, \boldsymbol{p}_n}) = w_{n,0}^B.
\end{aligned}$$

Remark 3.3. One can easily write the continued fraction expansion for the Galois conjugate $-\tilde{w}_{n,i}^B$ of $-w_{n,i}^B$ in terms of that of $w_{n,i}^B$. Indeed, let $(d_0, \overline{d_1, \dots, d_m})$ be the continued fraction expansion of $w_{n,i}^B$. The quadratic $ST^{-d_0}(w_{n,i}^B)$ is purely periodic with continued fraction $(\overline{d_1, \dots, d_m})$ so, by the property (III), its Galois conjugate is $1/(\overline{d_m, \dots, d_1})$. Therefore,

$$\tilde{w}_{n,i}^B = T^{d_0} S(1/(\overline{d_m, \dots, d_1})) = -(d_m - d_0, \overline{d_{m-1}, \dots, d_1, d_m}).$$

4. Convergence property

In this section we study the values of a modular function on the Markov tree. Let B be any branch of the tree and w_n^B be the n-th Markov quadratic on B. Let $A_{w_n^B} = A_0^{-1} \cdots A_{\ell_n}^{-1}$, where $A_0 = I$ and $A_i \in \{T^{-1}, V^{-1}\}$ for $1 \le i \le \ell_n$. Let $\rho = e^{\pi i/3}$ and $z_i = A_0^{-1} \cdots A_i^{-1} \rho^2$. Then using the modularity of f we have

$$\begin{aligned}
f(w_n^B) &= -\sqrt{D} \sum_{i=0}^{\ell_n - 1} \int_{z_i}^{z_{i+1}} \frac{f(z)}{Q_{w_n^B}(z, 1)}\, dz\\
&= -\sqrt{D} \sum_{i=0}^{\ell_n - 1} \int_{\rho^2}^{A_{i+1}^{-1}\rho^2} \frac{f(z)}{(Q_{w_n^B} | A_0^{-1} \cdots A_i^{-1})(z, 1)}\, dz\\
&= -\sum_{i=0}^{\ell_n - 1} \int_{\rho^2}^{A_{i+1}^{-1}\rho^2} f(z) \left(\frac{1}{z - w_{n,i}^B} - \frac{1}{z - \tilde{w}_{n,i}^B} \right) dz.
\end{aligned}$$

Since $V(\rho^2) = T(\rho^2) = \rho$, we obtain:

Lemma 4.1. *For $n \ge 0$ we have*

$$f(w_n^B) = \int_{\rho}^{\rho^2} \sum_{i=0}^{\ell_n - 1} f(z) \left(\frac{1}{z - w_{n,i}^B} - \frac{1}{z - \tilde{w}_{n,i}^B} \right) dz. \tag{12}$$

Lemma 4.1 is the main tool we use to estimate the values of modular functions at real quadratic irrationalities.

Throughout the paper, we denote by $\mathcal{C}$ the arc of circle joining ρ^2 and ρ. We denote by ε_n^B the image of $A_{w_n}^B$ under the isomorphism (2), so the length of $C_{w_n^B}$ equals $2\log\varepsilon_n^B$.

Our first goal is to show that the normalized values $f^{\mathrm{nor}}(w_n^B)$ for any modular function f along any branch B converge to the value $f^{\mathrm{nor}}(w_0^B)$. We call this property "convergence property" and prove it in this section. The main idea of the proof is to divide the sum in Lemma 4.1 into several ranges and bound each piece making repeated use of Lemma 2.1. For simplicity of the notation, as mentioned before, we restrict to a left but not the leftmost branch. However, the argument in the proof of Theorem 4.2 applies in the same way if B is a right branch or $B = L$. Only the bound $\delta_1(r, N)$ will be slightly modified but will still be of the form $O(rN\lambda^{rN})$ where $\lambda = (2/(1+\sqrt{5}))^2$. Hence Corollary 4.4 also remains true for any branch.

Theorem 4.2. *Let f be a modular function, B be any left branch $\neq L$ of the Markov tree $\mathcal{T}$ and $N \geq 1$. There exists a complex number $K = K_{f,B,N}$ such that for all $n \geq N$,*

$$\left| f(w_n^B) - nf(w_0^B) - K \right| \leq \delta_1(r, N) \max_{z\in\mathcal{C}} |f(z)|, \tag{13}$$

where

$$\delta_1(r, N) = \frac{80\pi}{3}(2 + r(N+1))\left(\frac{2}{1+\sqrt{5}}\right)^{2(rN-1)} \tag{14}$$

and $r+1$ is the number of partial quotients in the period of w_0^B.

Proof. By applying Lemma 4.1 for $f(w_n^B)$ and $f(w_0^B)$ we have:

$$f(w_n^B) - nf(w_0^B) = \int_\rho^{\rho^2} f(z)(S_1(n, N, z) + S_2(n, N, z) + S_3(n, N, z))\, dz, \tag{15}$$

where

$$S_1(n, N, z) = \sum_{i=0}^{a-1} \frac{1}{z - w_{n,i}^B} + \sum_{i=a+(n-N)\ell_0}^{\ell_n - 1} \frac{1}{z - w_{n,i}^B} - \sum_{i=0}^{a+N\ell_0-1} \frac{1}{z - \tilde{w}_{n,i}^B} - N\sum_{i=0}^{\ell_0-1}\left(\frac{1}{z - w_{0,i}^B} - \frac{1}{z - \tilde{w}_{0,i}^B}\right),$$

$$S_2(n, N, z) = \sum_{i=a}^{a+(n-N)\ell_0-1} \frac{1}{z - w_{n,i}^B} - (n-N)\sum_{i=0}^{\ell_0-1} \frac{1}{z - w_{0,i}^B},$$

$$S_3(n, N, z) = -\sum_{i=a+N\ell_0}^{a+n\ell_0-1} \frac{1}{z - \tilde{w}_{n,i}^B} + (n-N)\sum_{i=0}^{\ell_0-1} \frac{1}{z - \tilde{w}_{0,i}^B}.$$

Moreover, we can also write

$$S_1(n, N, z) = S_1(N, N, z) + (S_1(n, N, z) - S_1(N, N, z)). \tag{16}$$

Define

$$K := \int_{\rho}^{\rho^2} f(z) S_1(N, N, z)\, dz$$

and

$$c(n, z) := |S_1(n, N, z) - S_1(N, N, z)| + |S_2(n, N, z)| + |S_3(n, N, z)|.$$

Then

$$|f(w_n^B) - nf(w_0^B) - K| \le \int_{\rho}^{\rho^2} c(n, z)|f(z)|\,|dz|. \tag{17}$$

These divisions are guided by the continued fraction expansions of all the terms in the cycle of w_n^B and w_0^B and their conjugates. As we will see shortly, the repeated use of Lemma 2.1 will allow us to bound all the other sums after we separate the main term K.

Let $\lambda = (2/(1+\sqrt{5}))^2$. If we can show that

$$c(n, z) \le 80(2 + r(N+1))\lambda^{rN-1} \tag{18}$$

for $z \in \mathcal{C}$, then the theorem is proved. Next we show (18).

Bound for $|S_2(n, N, z)|$**.** We have that

$$|S_2(n, N, z)| \le \sum_{k=0}^{n-N-2} \sum_{i=1}^{\ell_0} \frac{|w^B_{n,2+a+k\ell_0+i} - w^B_{0,2+i}|}{|z - w^B_{n,2+a+k\ell_0+i}||z - w^B_{0,2+i}|} + \sum_{i=0}^{2} \frac{|w^B_{n,a+i} - w^B_{0,i}|}{|z - w^B_{n,a+i}||z - w^B_{0,i}|} + \sum_{i=1}^{\ell_0-3} \frac{|w^B_{n,2+a+(n-N-1)\ell_0+i} - w^B_{0,2+i}|}{|z - w^B_{n,2+a+(n-N-1)\ell_0+i}||z - w^B_{0,2+i}|}.$$

Clearly for any $z \in \mathcal{C}$ and $x \in \mathbb{R}$, we have that $|z - x| \ge \mathrm{Im}(e^{2\pi i/3}) = \sqrt{3/2}$. Hence the denominators are bounded below by $\frac{3}{4}$ when $z \in \mathcal{C}$ since the points w are real. The numerators can be bounded by using Lemma 2.1. For $i = 0, 1, 2$,

$$w^B_{n,a+i} = (3 - i, \overline{\boldsymbol{p}_n, a_1, \ldots, a_s}) \quad \text{and} \quad w^B_{0,i} = (3 - i, \overline{\boldsymbol{p}})$$

coincide at least in the first $rn+1$ partial quotients. For each $0 \le k \le n-N-2$, we have: For $1 \le i \le b_1 - 1$,

$$w^B_{n,2+a+k\ell_0+i} = \big(b_1 - i, \overline{b_2, \ldots, b_r, \boldsymbol{p}_{n-1-k}, a_1, \ldots, a_s, \boldsymbol{p}_k, b_1}\big). \tag{19}$$

For the next $b_2 - 1$ values of i $(b_1 \le i \le b_1 + b_2 - 2)$,

$$w^B_{n,2+a+k\ell_0+i} = \big(b_2 - j, \overline{b_3, \ldots, b_r, \boldsymbol{p}_{n-1-k}, a_1, \ldots, a_s, \boldsymbol{p}_k, b_1, b_2}\big) \tag{20}$$

with $1 \le j \le b_2 - 1$. This process goes on until the last $b_r - 1$ values of i, where

$$w^B_{n,2+a+k\ell_0+i} = \big(b_r - j, \overline{\boldsymbol{p}_{n-1-k}, a_1, \ldots, a_s, \boldsymbol{p}_{k+1}}\big)$$

with $1 \le j \le b_r - 1$. For $k = n - N - 1$, we have the same pattern as before except for the last block of values of i, where we only have $b_r - 3$ of them.

Now, for each $0 \le k \le n-N-1$, for $1 \le i \le b_1 - 1$, (19) and

$$w^B_{0,2+i} = (b_1 - i, \overline{b_2, \dots, b_r, b_1})$$

coincide in the first $rn - rk$ partial quotients. For the next $b_2 - 1$ values of i, (20) and

$$w^B_{0,2+i} = (b_2 - j, \overline{b_3, \dots, b_r, b_1, b_2}) \qquad (1 \le j \le b_2 - 1)$$

coincide in the first $rn - rk - 1$ partial quotients, similarly for the next $b_3 - 1$ values of i, $w^B_{n,2+a+k\ell_0+i}$ and $w^B_{0,2+i}$ coincide in the first $rn - rk - 2$ partial quotients, etc. Therefore, using Lemma 2.1, for $z \in \mathcal{C}$, we have

$$\begin{aligned}
|S_2(n,N,z)| &\le \frac{40}{3}\left(3\lambda^{rn} + \sum_{i=1}^{r}(b_i - 1)\sum_{k=0}^{n-N-1}\lambda^{r(n-k)-i}\right)\\
&\le \frac{40}{3}\left(3\lambda^{rn} + 3\left(\sum_{i=1}^{r}\lambda^{-i}\right)\left(\sum_{k=N+1}^{n}\lambda^{rk}\right)\right)\\
&\le \frac{40}{3}\left(3\lambda^{rn} + 3\left(\sum_{i=1}^{r}\lambda^{r-i}\right)\left(\sum_{k=N}^{n-1}\lambda^{rk}\right)\right)\\
&\le \frac{40}{3}\left(3\lambda^{rn} + 3\left(\sum_{i=0}^{r-1}\lambda^{i}\right)\left(\sum_{k=N}^{n-1}\lambda^{rk}\right)\right)\\
&\le 40\lambda^{rN}\left(1 + \frac{1}{(1-\lambda)(1-\lambda^r)}\right)\\
&\le 120\lambda^{rN}. \qquad (21)
\end{aligned}$$

In the second inequality we used that $b_i \le 4$, whereas the last inequality follows from the numerical value $1/1-\lambda = 1.618\dots$

Bound for $|S_3(n,N,z)|$**.** In a similar way we bound $|S_3(n,N,z)|$. We have that

$$|S_3(n,N,z)| \le \sum_{k=N}^{n-2}\sum_{i=1}^{\ell_0}\frac{|\tilde{w}^B_{n,2+a+k\ell_0+i} - \tilde{w}^B_{0,2+i}|}{|z - \tilde{w}^B_{n,2+a+k\ell_0+i}||z - \tilde{w}^B_{0,2+i}|} + \sum_{i=0}^{2}\frac{|\tilde{w}^B_{n,a+N\ell_0+i} - \tilde{w}^B_{0,i}|}{|z - \tilde{w}^B_{n,a+N\ell_0+i}||z - \tilde{w}^B_{0,i}|} + \sum_{i=1}^{\ell_0-3}\frac{|\tilde{w}^B_{n,2+a+(n-1)\ell_0+i} - \tilde{w}^B_{0,2+i}|}{|z - \tilde{w}^B_{n,2+a+(n-1)\ell_0+i}||z - \tilde{w}^B_{0,2+i}|}.$$

For $i = 0, 1, 2$, using Remark 3.3, we have that

$$-\tilde{w}^B_{n,a+N\ell_0+i} = \left(1 + i, \overline{b_{r-1}, \dots, b_1, \boldsymbol{q}_{N-1}, a_s, \dots, a_1, q_{n-N}, b_r}\right)$$

and

$$-\tilde{w}^B_{0,i} = (1 + i, \overline{b_{r-1}, \dots, b_1, b_r})$$

coincide in the first rN partial quotients. For each $N \le k \le n-2$, we have: For $1 \le i \le b_1 - 1$,

$$-\tilde{w}^B_{n,2+a+k\ell_0+i} = (i, \overline{\boldsymbol{q}_k, a_s \ldots, a_1, \boldsymbol{q}_{n-k}}). \tag{22}$$

For the next $b_2 - 1$ values of i $(b_1 \le i \le b_1 + b_2 - 2)$,

$$-\tilde{w}^B_{n,2+a+k\ell_0+i} = \big(j, \overline{b_1, \boldsymbol{q}_k, a_s \ldots, a_1, \boldsymbol{q}_{n-1-k}, b_r, \ldots, b_3, b_2}\big) \tag{23}$$

with $1 \le j \le b_2 - 1$. This process goes on until the last $b_r - 1$ values of i, where

$$-\tilde{w}^B_{n,2+a+k\ell_0+i} = \big(j, \overline{b_{r-1}, \ldots, b_1, \boldsymbol{q}_k, a_s, \ldots, a_1, \boldsymbol{q}_{n-1-k}, b_r}\big)$$

with $1 \le j \le b_r - 1$.

For $k = n-1$, we have the same pattern as before except for the last block of values of i, where we only have $b_r - 3$ of them. Now, for each $N \le k \le n-1$, for the first $b_1 - 1$ values of i, (22) coincide with

$$-\tilde{w}^B_{0,2+i} = (i, \overline{\boldsymbol{q}})$$

in the first $rk+1$ partial quotients. For the next $b_2 - 1$ values of i, (23) coincide with

$$-\tilde{w}^B_{0,2+i} = (j, \overline{b_1, b_r, \ldots, b_2}) \quad (1 \le j \le b_2 - 1)$$

in the first $rk+2$ partial quotients, for the next $b_3 - 1$ i-values, $-\tilde{w}^B_{n,2+a+k\ell_0+i}$ and $-\tilde{w}^B_{0,2+i}$ coincide in the first $rk+3$ partial quotients, etc. Once again using Lemma 2.1, and the fact that $b_i \le 4$ together with the numerical value of λ, we have, for $z \in \mathcal{C}$,

$$\begin{aligned}
|S_3(n, N, z)| &\le \frac{40}{3}\left(3\lambda^{rN-1} + \sum_{i=1}^{r}(b_i - 1)\sum_{k=N}^{n-1}\lambda^{rk+i-1}\right) \\
&\le \frac{40}{3}\left(3\lambda^{rN-1} + 3\left(\sum_{i=1}^{r}\lambda^{i-1}\right)\left(\sum_{k=N}^{n-1}\lambda^{rk}\right)\right) \\
&\le 40\lambda^{rN-1}\left(1 + \frac{\lambda}{(1-\lambda)(1-\lambda^r)}\right) \\
&\le 80\lambda^{rN-1}.
\end{aligned} \tag{24}$$

Bound for $|S_1(n, N, z) - S_1(N, N, z)|$. We have

$$\begin{aligned}
&|S_1(n, N, z) - S_1(N, N, z)| \\
&\quad\le \sum_{i=0}^{a-1}\frac{|w^B_{n,i} - w^B_{N,i}|}{|z - w^B_{n,i}||z - w^B_{N,i}|} + \sum_{i=a}^{\ell_N - 1}\frac{|w^B_{n,i+(n-N)\ell_0} - w^B_{N,i}|}{|z - w^B_{n,i+(n-N)\ell_0}||z - w^B_{N,i}|} + \sum_{i=0}^{a+N\ell_0-1}\frac{|\tilde{w}^B_{n,i} - \tilde{w}^B_{N,i}|}{|z - \tilde{w}^B_{n,i}||z - \tilde{w}^B_{N,i}|}.
\end{aligned}$$

Again the denominators are bounded below by $\frac{3}{4}$ for $z \in \mathcal{C}$ and we use Lemma 2.1 to bound the numerators. For the first term in the first sum, using

$$w^B_{n,0} = (3, \overline{a_1, \ldots, a_s, \boldsymbol{p}_n}) \tag{25}$$

and

$$w^B_{N,0} = (3, \overline{a_1, \dots, a_s, \boldsymbol{p}_N}), \tag{26}$$

one can see that the successive terms $w^B_{n,i}$ and $w^B_{N,i}$ (up to $i = a-1$) coincide at least in the first rN partial quotients. This is also true for the second sum, where we have

$$w^B_{n,a+(n-N)\ell_0} = (3, \overline{\boldsymbol{p}_{n-N}, a_1, \dots, a_s, \boldsymbol{p}_N}) \quad \text{and} \quad w^B_{N,a} = (3, \overline{\boldsymbol{p}_N, a_1, \dots, a_s}),$$

as well as for the third and fourth sums, where we can use Remark 3.3 and the continued fractions of (25) and (26), and

$$w^B_{n,a+n\ell_0} = (3, \overline{a_1, \dots, a_s, \boldsymbol{p}_n}), \quad \text{and} \quad w^B_{n,a+N\ell_0} = (3, \overline{a_1, \dots, a_s, \boldsymbol{p}_N}),$$

respectively. Hence, using (11), we have

$$|S_1(n, N, z) - S_1(N, N, z)| \le \tfrac{80}{3}\ell_N \lambda^{rN-1} \overset{(11)}{\le} 80r(N+1)\lambda^{rN-1}. \tag{27}$$

Finally, since $\lambda < \frac{2}{3}$, the bounds (21), (24) and (27) give

$$c(n, z) \le 80(2 + r(N+1))\lambda^{rN-1}. \qquad \square$$

In particular, Theorem 4.2 applied to the function $f = 1$ gives:

Corollary 4.3. *Let B be any left branch $\neq L$ of $\mathcal{T}$ and $N \ge 1$. For all $n \ge N$, there exists $K = K_{B,N} \in \mathbb{R}$ such that*

$$|\log \varepsilon^B_n - n \log \varepsilon^B_0 - K| \le \delta_1(r, N) \tag{28}$$

with $\delta_1(r, N)$ and r as in (14).

The next corollary proves Theorem 1.1 from the introduction.

Corollary 4.4. *Let f be a modular function. For any left branch $B \neq L$ of $\mathcal{T}$,*

$$\lim_{n\to\infty} f^{\mathrm{nor}}(w^B_n) = f^{\mathrm{nor}}(w^B_0).$$

Proof. It follows from Theorem 4.2 and Corollary 4.3 that $|f(w^B_n) - nf(w^B_0)|$ and $|\log \varepsilon^B_n - n \log \varepsilon^B_0|$ are bounded above and below by absolute constants (not depending on n). Then

$$0 = \lim_{n\to\infty} \frac{|f(w^B_n) - nf(w^B_0)|}{\log \varepsilon^B_n} = \lim_{n\to\infty} \left| \frac{f(w^B_n)}{\log \varepsilon^B_n} - \frac{f(w^B_0)}{\log \varepsilon^B_0} \right|. \qquad \square$$

5. Interlacing property

In this section we prove Theorem 1.2. As in the proof of the convergence property we restrict again to a left but not the leftmost branch in what follows.The argument applies in the same way to any branch, with the bound $\delta_2(n, r)$ slightly modified. It will still be of the form $O(rn\lambda^{rn})$. Hence Theorem 1.2 applies in fact to any branch of the Markov tree and it is a consequence of the next theorem whose proof is similar to the proof of Theorem 4.2.

Theorem 5.1. *Let f be a modular function. For every left branch $B \neq L$ of the Markov tree $\mathcal{T}$ and for all $n \geq 1$,*

$$\left| f(w_{n+1}^B) - f(w_n^B) - f(w_0^B) \right| \leq \delta_2(n, r) \max_{z \in \mathcal{C}} |f(z)| \tag{29}$$

where

$$\delta_2(n, r) = \frac{80\pi}{3}(n+2)r\left(\frac{2}{1+\sqrt{5}}\right)^{2(rn-1)} \tag{30}$$

and $r+1$ is the number of partial quotients in the period of w_0^B.

Proof. Once again, applying Lemma 4.1 and (10) gives

$$f(w_{n+1}^B) - f(w_n^B) = f(w_0^B) + \int_\rho^{\rho^2} f(z) R_1(n, z)\, dz + \int_\rho^{\rho^2} f(z) R_2(n, z)\, dz, \tag{31}$$

where

$$R_1(n, z) = \sum_{i=0}^{a-1}\left(\frac{1}{z - w_{n+1,i}^B} - \frac{1}{z - w_{n,i}^B}\right) + \sum_{i=a}^{\ell_n - 1}\left(\frac{1}{z - w_{n+1,\ell_0+i}^B} - \frac{1}{z - w_{n,i}^B}\right) - \sum_{i=0}^{\ell_n - 1}\left(\frac{1}{z - \tilde{w}_{n+1,i}^B} - \frac{1}{z - \tilde{w}_{n,i}^B}\right),$$

$$R_2(n, z) = \sum_{i=0}^{\ell_0 - 1}\left(\frac{1}{z - w_{n+1,a+i}^B} - \frac{1}{z - w_{0,i}^B} - \frac{1}{z - \tilde{w}_{n+1,\ell_n+i}^B} + \frac{1}{z - \tilde{w}_{0,i}^B}\right).$$

Next we give upper bounds for the norms of the two sums above when $z \in \mathcal{C}$. We set again $\lambda = (2/(1+\sqrt{5}))^2$.

Bound for $|R_1(n, z)|$**.** For $z \in \mathcal{C}$, we have

$$|R_1(n, z)| \leq \sum_{i=0}^{a-1} \frac{|w_{n+1,i}^B - w_{n,i}^B|}{|z - w_{n+1,i}^B||z - w_{n,i}^B|} + \sum_{i=a}^{\ell_n - 1} \frac{|w_{n+1,i+\ell_0}^B - w_{n,i}^B|}{|z - w_{n+1,i+\ell_0}^B||z - w_{n,i}^B|} + \sum_{i=0}^{a+n\ell_0-1} \frac{|\tilde{w}_{n+1,i}^B - \tilde{w}_{n,i}^B|}{|z - \tilde{w}_{n+1,i}^B||z - \tilde{w}_{n,i}^B|} + \sum_{i=a+n\ell_0}^{\ell_n - 1} \frac{|\tilde{w}_{n+1,i}^B - \tilde{w}_{n,i}^B|}{|z - \tilde{w}_{n+1,i}^B||z - \tilde{w}_{n,i}^B|}.$$

As before we use the bound of $\frac{3}{4}$ for the denominators and Lemma 2.1 for the numerators. In the first sum using

$$w_{n+1,0}^B = (3, \overline{a_1, \ldots, a_s, \boldsymbol{p}_{n+1}}) \tag{32}$$

and

$$w_{n,0}^B = (3, \overline{a_1, \ldots, a_s, \boldsymbol{p}_n}), \tag{33}$$

one can see that the successive terms $w_{n+1,i}^B$ and $w_{n,i}^B$ (up to $i = a-1$) coincide at least in the first rn partial quotients. The same is true for the second sum, where

$$w_{n+1,a+\ell_0}^B = (3, \overline{\boldsymbol{p}_n, a_1, \ldots, a_s, \boldsymbol{p}}) \quad \text{and} \quad w_{n,a}^B = (3, \overline{\boldsymbol{p}_n, a_1, \ldots, a_s}).$$

For the third and fourth sums, we use once again Remark 3.3 together with (32) and (33), and

$$w^B_{n+1,a+n\ell_0} = (3, \overline{a_1, \dots, a_s, \boldsymbol{p}_{n+1}}) \quad \text{and} \quad w^B_{n,a+n\ell_0} = (3, \overline{a_1, \dots, a_s, \boldsymbol{p}_n})$$

respectively.

Hence

$$|R_1(n,z)| \le \tfrac{80}{3}\ell_n \lambda^{rn-1} \overset{(11)}{\le} 80r(n+1)\lambda^{rn-1}.$$

Bound for $|R_2(n,z)|$**.** In a similar way we bound this second sum when $z \in \mathcal{C}$:

$$|R_2(n,z)| \le \sum_{i=0}^{\ell_0-1} \frac{|w^B_{n+1,a+i} - w^B_{0,i}|}{|z - w^B_{n+1,a+i}||z - w^B_{0,i}|} + \frac{|\tilde{w}^B_{n+1,\ell_n+i} - \tilde{w}^B_{0,i}|}{|z - \tilde{w}^B_{n+1,\ell_n+i}||z - \tilde{w}^B_{0,i}|}.$$

Again using

$$w^B_{n+1,a} = (3, \overline{\boldsymbol{p}_{n+1}})$$

and

$$w^B_{0,0} = (3, \bar{\boldsymbol{p}}), \tag{34}$$

one can see that all the successive terms $w^B_{n+1,a+i}$ and $w^B_{0,i}$ in the sum coincide at least in the first rn partial quotients. For the conjugate terms, one can see from (34) and

$$w^B_{n+1,\ell_n} = (3, \overline{\boldsymbol{p}, a_1, \dots, a_s, \boldsymbol{p}_n})$$

that $-\tilde{w}^B_{n+1,\ell_n+i}$ and $-\tilde{w}^B_{0,i}$ coincide as well in the first rn partial quotients. Hence

$$|R_2(n,z)| \le \tfrac{80}{3}\ell_0 \lambda^{rn-1} \overset{(11)}{\le} 80r\lambda^{rn-1}.$$

Therefore,

$$|f(w_{n+1}) - f(w_n) - f(w_0)| \le \int_\rho^{\rho^2} |f(z)|(|R_1(n,z)| + |R_2(n,z)|)\,|dz| \le \delta_2(n,r) \max_{z\in\mathcal{C}} |f(z)|$$

with

$$\delta_2(n,r) = \frac{80\pi}{3} r(n+2)\lambda^{rn-1}. \qquad \square$$

Theorem 5.1 applied to the function $f = 1$ gives:

Corollary 5.2. *For every left branch $B \neq L$ of $\mathcal{T}$ and for all $n \ge 1$,*

$$\left|\log \varepsilon^B_{n+1} - \log \varepsilon^B_n - \log \varepsilon^B_0\right| \le \delta_2(n,r)$$

with $\delta_2(n,r)$ and r as in (30).

We finish this section by giving the proof of Theorem 1.2 in the case that the branch B is any left branch $\neq L$. The proof of the general case goes along the same lines.

Theorem 5.3. *Let f be a modular function, B be any left branch $\neq L$ of the Markov tree $\mathcal{T}$. There exists a constant $N_{f,B}$ such that, for all $n \ge N_{f,B}$, the real and imaginary parts of $f^{\mathrm{nor}}(w^B_{n+1})$ lie between the real and imaginary parts respectively of $f^{\mathrm{nor}}(w^B_0)$ and $f^{\mathrm{nor}}(w^B_n)$.*

Proof. By definition, the inequality

$$\mathrm{Re}(f^{\mathrm{nor}}(w_n^B)) < \mathrm{Re}(f^{\mathrm{nor}}(w_0^B))$$

holds if and only if

$$\mathrm{Re}(f(w_n^B)) \log \varepsilon_0^B < \mathrm{Re}(f(w_0^B)) \log \varepsilon_n^B. \tag{35}$$

Let N, M be positive constants. For all $n \geq \max(N, M)$, we can write

$$\mathrm{Re}(f(w_n^B)) = n\,\mathrm{Re}(f(w_0^B)) + K_{f,B,N} + \varepsilon_1(n, N), \tag{36}$$

$$\log \varepsilon_n^B = n \log \varepsilon_0^B + K_{1,B,M} + \varepsilon_2(n, M), \tag{37}$$

where $K_{f,B,N}$, $K_{1,B,M}$ are the real parts of the constants in Theorem 4.2 and Corollary 4.3 respectively, $|\varepsilon_1(n, N)| \leq \delta_1(N) \max_{z\in\mathcal{C}} |f(z)|$ and $|\varepsilon_2(n, M)| \leq \delta_1(M)$. Therefore (35) is equivalent to

$$\mathrm{Re}(f^{\mathrm{nor}}(w_0^B)) > \frac{K_{f,B,N}}{K_{1,B,M}} + \frac{\varepsilon_1(n, N) - \varepsilon_2(n, M)\,\mathrm{Re}(f^{\mathrm{nor}}(w_0^B))}{K_{1,B,M}}. \tag{38}$$

There exists a constant $C_1(f, B)$ depending on f and B such that, for $\max(N, M) \geq C_1(f, B)$, (38) is equivalent to either

$$\mathrm{Re}(f^{\mathrm{nor}}(w_0^B)) > \frac{K_{f,B,N}}{K_{1,B,M}} \tag{39}$$

or (39) with the strict inequality replaced by $\geq$, according to whether the error term in (38) is positive or negative. If we can choose $N, M \geq C_1(f, B)$ satisfying $\mathrm{Re}(f^{\mathrm{nor}}(w_0^B)) \neq K_{f,B,N}/K_{1,B,M}$, then (38) is equivalent to (39) for those N, M. If we cannot choose such N, M, then $K_{f,B,N}$, $K_{1,B,M}$ would be constants that do not depend on N, M, and in particular $\varepsilon_1(n, N) = \varepsilon_2(n, M) = 0$. Hence, also in this case (38) is equivalent to (39) for all $N, M \geq C_1(f, B)$.

In a similar way, the inequality

$$\mathrm{Re}(f^{\mathrm{nor}}(w_n^B)) > \mathrm{Re}(f^{\mathrm{nor}}(w_0^B))$$

is equivalent to

$$\mathrm{Re}(f^{\mathrm{nor}}(w_0^B)) < \frac{K_{f,B,N}}{K_{1,B,M}} \tag{40}$$

for N, M chosen as before. Since (39) and (40) do not depend on n, we have either

$$\mathrm{Re}(f^{\mathrm{nor}}(w_n^B)) < \mathrm{Re}(f^{\mathrm{nor}}(w_0^B))$$

simultaneously for all $n \geq \max(N, M)$ with N, M chosen as before, or

$$\mathrm{Re}(f^{\mathrm{nor}}(w_n^B)) > \mathrm{Re}(f^{\mathrm{nor}}(w_0^B)).$$

Similarly, the inequality

$$\mathrm{Re}(f^{\mathrm{nor}}(w_{n+1}^B)) < \mathrm{Re}(f^{\mathrm{nor}}(w_n^B))$$

holds if and only if

$$\mathrm{Re}(f(w_{n+1}^B)) \log \varepsilon_n^B < \mathrm{Re}(f(w_n^B)) \log \varepsilon_{n+1}^B. \tag{41}$$

Theorem 5.1 and Corollary 5.2 respectively imply that

$$\operatorname{Re}(f(w_{n+1}^B)) = \operatorname{Re}(f(w_n^B)) + \operatorname{Re}(f(w_0^B)) + \mu(n)$$

with $|\mu(n)| \le \delta_2(n) \max_{z\in\mathcal{C}} |f(z)|$ and

$$\log \varepsilon_{n+1}^B = \log \varepsilon_n^B + \log \varepsilon_0^B + \nu(n)$$

with $|\nu(n)| \le \delta_2(n)$. Hence (41) is equivalent to

$$(\operatorname{Re}(f(w_0^B)) + \mu(n)) \log \varepsilon_n^B < \operatorname{Re}(f(w_n^B))(\log \varepsilon_0^B + \nu(n)). \tag{42}$$

Now, there exists a constant $C_2(f, B) \ge C_1(f, B)$ such that, for $n \ge C_2(f, B)$, we have that

$$\operatorname{Re}(f^{\mathrm{nor}}(w_n^B)) \ne \operatorname{Re}(f^{\mathrm{nor}}(w_0^B))$$

and that (42) is equivalent to

$$\operatorname{Re}(f(w_0^B)) \log \varepsilon_n^B < \operatorname{Re}(f(w_n^B)) \log \varepsilon_0^B. \tag{43}$$

Using (36) and (37) again, we obtain that (43) is equivalent to

$$\operatorname{Re}(f^{\mathrm{nor}}(w_0^B)) < \frac{K_{f,B,N}}{K_{1,B,M}}, \tag{44}$$

where N, M are chosen as before.

Therefore, we finally have that either

$$\operatorname{Re}(f^{\mathrm{nor}}(w_0^B)) < \operatorname{Re}(f^{\mathrm{nor}}(w_{n+1}^B)) < \operatorname{Re}(f^{\mathrm{nor}}(w_n^B))$$

for all $n \ge \max(C_2(f, B), N, M)$ or

$$\operatorname{Re}(f^{\mathrm{nor}}(w_n^B)) < \operatorname{Re}(f^{\mathrm{nor}}(w_{n+1}^B)) < \operatorname{Re}(f^{\mathrm{nor}}(w_0^B)).$$

The same argument applies to the imaginary parts of $f^{\mathrm{nor}}(w_{n+1}^B)$, $f^{\mathrm{nor}}(w_n^B)$ and $f^{\mathrm{nor}}(w_0^B)$. □

Acknowledgements

We thank M. Kaneko and the referee for numerous and very helpful comments that improved our exposition.

References

[Aigner 2013] M. Aigner, *Markov's theorem and 100 years of the uniqueness conjecture*, Springer, 2013. MR Zbl

[Bombieri 2007] E. Bombieri, "Continued fractions and the Markoff tree", *Expo. Math.* **25**:3 (2007), 187–213. MR Zbl

[Bourgain et al. 2016a] J. Bourgain, A. Gamburd, and P. Sarnak, "Markoff surfaces and strong approximation, I", preprint, 2016. arXiv

[Bourgain et al. 2016b] J. Bourgain, A. Gamburd, and P. Sarnak, "Markoff triples and strong approximation", *C. R. Math. Acad. Sci. Paris* **354**:2 (2016), 131–135. MR

[Duke et al. 2011] W. Duke, O. Imamoḡlu, and A. Tóth, "Cycle integrals of the j-function and mock modular forms", *Ann. of Math.* (2) **173**:2 (2011), 947–981. MR Zbl

[Ghosh and Sarnak 2017] A. Ghosh and P. Sarnak, "Integral points on Markoff type cubic surfaces", preprint, 2017. arXiv

[Hall 1947] M. Hall, Jr., "On the sum and product of continued fractions", *Ann. of Math.* (2) **48** (1947), 966–993. MR Zbl

[Kaneko 2009] M. Kaneko, "Observations on the "values" of the elliptic modular function $j(\tau)$ at real quadratics", *Kyushu J. Math.* **63**:2 (2009), 353–364. MR Zbl

[Malyshev 1977] A. V. Malyshev, "Markov and Lagrange spectra (a survey of the literature)", pp. 5–38 in *Studies in number theory, part 4*, vol. 67, "Nauka" Leningrad. Otdel. Mat. Inst. Steklov. (LOMI), 1977. In Russian; translated in *J. Soviet Math.* **16**:1 (1981), 767–788. MR

[Markov 1879] A. Markoff, "Sur les formes quadratiques binaires indéfinies", *Math. Ann.* **15** (1879), 381–406. Zbl

[Markov 1880] A. Markoff, "Sur les formes quadratiques binaires indéfinies", *Math. Ann.* **17**:3 (1880), 379–399. MR Zbl

[Roth 1955] K. F. Roth, "Rational approximations to algebraic numbers", *Mathematika* **2** (1955), 1–20; corrigendum, 168. MR Zbl

[Zagier 1981] D. B. Zagier, *Zetafunktionen und quadratische Körper: eine Einführung in die höhere Zahlentheorie*, Springer, 1981. MR Zbl

Communicated by Philippe Michel
Received 2018-03-27 Revised 2018-12-18 Accepted 2019-02-08

paloma.bengoechea@math.ethz.ch *Department of Mathematics, ETH Zurich, Switzerland*

ozlem.imamoglu@math.ethz.ch *Department of Mathematics, ETH Zurich, Switzerland*

msp

ALGEBRA AND NUMBER THEORY 13:4 (2019)
dx.doi.org/10.2140/ant.2019.13.963

A finiteness theorem for specializations of dynatomic polynomials

David Krumm

Let t and x be indeterminates, let $\phi(x) = x^2 + t \in \mathbb{Q}(t)[x]$, and for every positive integer n let $\Phi_n(t, x)$ denote the n-th dynatomic polynomial of ϕ. Let G_n be the Galois group of Φ_n over the function field $\mathbb{Q}(t)$, and for $c \in \mathbb{Q}$ let $G_{n,c}$ be the Galois group of the specialized polynomial $\Phi_n(c, x)$. It follows from Hilbert's irreducibility theorem that for fixed n we have $G_n \cong G_{n,c}$ for every c outside a thin set $E_n \subset \mathbb{Q}$. By earlier work of Morton (for $n = 3$) and the present author (for $n = 4$), it is known that E_n is infinite if $n \leq 4$. In contrast, we show here that E_n is finite if $n \in \{5, 6, 7, 9\}$. As an application of this result we show that, for these values of n, the following holds with at most finitely many exceptions: for every $c \in \mathbb{Q}$, more than 81% of prime numbers p have the property that the polynomial $x^2 + c$ does not have a point of period n in the p-adic field $\mathbb{Q}_p$.

1. Introduction

Let c be a rational number and let $\phi_c(x) = x^2 + c$. Given any algebraic number x_0, we may consider the sequence $x_0, \phi_c(x_0), \phi_c(\phi_c(x_0)), \ldots$. If this sequence is periodic with period n, we say that x_0 has period n under iteration of ϕ_c. By allowing c and x_0 to vary in $\mathbb{Q}$, one can find examples where x_0 has period 1, 2, or 3 under ϕ_c. For instance, the pairs

$$(c, x_0) = (0, 0),\ (-1, 0),\ \left(\tfrac{-29}{16}, \tfrac{5}{4}\right)$$

provide examples of periods 1, 2, and 3, respectively.

Poonen [1998] conjectured that if $n > 3$, then there does not exist $c \in \mathbb{Q}$ such that the polynomial ϕ_c has a rational point of period n. This has been proved for periods 4 and 5, and also for period 6 assuming the Birch–Swinnerton-Dyer conjecture; see [Morton 1998; Flynn et al. 1997; Stoll 2008]. The present paper is concerned with a strong form of Poonen's conjecture which was stated by the author in [Krumm 2016]: if $n > 3$, then for every $c \in \mathbb{Q}$ there exist infinitely many primes p such that ϕ_c does not have a point of period n in the p-adic field $\mathbb{Q}_p$. In fact, we will consider here a further strengthening of Poonen's conjecture.

Conjecture 1.1. *Fix $n > 3$. For every $c \in \mathbb{Q}$, let $T_{n,c}$ denote the set of primes p such that ϕ_c does not have a point of period n in $\mathbb{Q}_p$, and let $\delta(T_{n,c})$ be the Dirichlet density of $T_{n,c}$. Then $\delta(T_{n,c}) > 0$ for all $c \in \mathbb{Q}$.*

MSC2010: primary 37P05; secondary 11S15, 37P35.
Keywords: arithmetic dynamics, function fields, Galois theory.

In order to study these conjectures it is useful to consider a family of *dynatomic polynomials* defined as follows. For every positive integer n we define a two-variable polynomial $\Phi_n \in \mathbb{Q}[t, x]$ by the formula

$$\Phi_n(t, x) = \prod_{d \mid n} (\phi^d(x) - x)^{\mu(n/d)}, \tag{1-1}$$

where μ is the Möbius function, $\phi(x) = x^2 + t \in \mathbb{Q}(t)[x]$, and ϕ^d denotes the d-fold composition of ϕ with itself. The key property linking Φ_n to the above conjectures is that, for fixed $c \in \mathbb{Q}$, every algebraic number having period n under iteration of ϕ_c is a root of $\Phi_n(c, x)$, and conversely, every root of $\Phi_n(c, x)$ has period n under ϕ_c except in rare cases when the period may be smaller than n; see [Morton and Patel 1994, Theorem 2.4] for further details.

Questions about the points of period n under ϕ_c can thus be phrased as questions about the roots of $\Phi_n(c, x)$. It is therefore to be expected that a good understanding of the Galois group of $\Phi_n(c, x)$ will yield substantial information about the dynamical properties of the map ϕ_c. The results of the article [Krumm 2018b] provide an example of the type of information that can be obtained in this way. By a careful analysis of how the Galois group of $\Phi_4(c, x)$ can change as c varies in $\mathbb{Q}$, it is proved there that if $\alpha \in \overline{\mathbb{Q}}$ has period four under a map ϕ_c, then the degree $[\mathbb{Q}(\alpha) : \mathbb{Q}]$ can only be 2, 4, 8, or 12; in particular the degree cannot be 1, which implies that ϕ_c does not have a rational point of period 4. Furthermore, the Galois group data is used to show that $\delta(T_{4,c}) > 0.39$ for every $c \in \mathbb{Q}$, thus proving Conjecture 1.1 for $n = 4$. Motivated by these results, we are led to the following problem.

Problem 1.2. Let $G_{n,c}$ denote the Galois group of $\Phi_n(c, x)$ over $\mathbb{Q}$. For fixed n, determine the structure of all the groups $G_{n,c}$ as c varies in $\mathbb{Q}$.

Since the polynomials $\Phi_n(c, x)$ for $c \in \mathbb{Q}$ are specializations of Φ_n, it follows from Hilbert's irreducibility theorem [Serre 2008, Proposition 3.3.5] that for every rational number c outside a thin subset of $\mathbb{Q}$, the group $G_{n,c}$ is isomorphic to the Galois group of Φ_n over the function field $\mathbb{Q}(t)$. Moreover, by work of Bousch [1992, Chapter 3] it is known that Φ_n is irreducible and that its Galois group, which we denote by G_n, is isomorphic to a wreath product of a cyclic group and a symmetric group; indeed, $G_n \cong (\mathbb{Z}/n\mathbb{Z}) \wr S_r$, where $rn = \deg \Phi_n$. Hence, for most $c \in \mathbb{Q}$ the structure of $G_{n,c}$ is known. However, a complete solution of Problem 1.2 would require understanding precisely for which numbers c the specialization $t \mapsto c$ fails to preserve the Galois group of Φ_n. This raises a new but closely related problem.

Problem 1.3. For fixed n, determine all $c \in \mathbb{Q}$ such that $G_{n,c} \ncong G_n$.

Let $E_n = \{c \in \mathbb{Q} \mid G_{n,c} \ncong G_n\}$. By work of Morton [1992] and the author [Krumm 2018b], the sets E_n are well understood for $n \leq 4$; in particular, one notable feature of these sets is that they are infinite. In contrast, empirical evidence suggests that E_n is finite for every $n > 4$. The main purpose of this article is to prove this finiteness statement for several values of n.

Theorem 1.4. *The set E_n is finite if $n \in \{5, 6, 7, 9\}$.*

Using this theorem we can provide further evidence in support of Conjecture 1.1. It follows from the theorem that, for the above values of n, we have $G_{n,c} \cong (\mathbb{Z}/n\mathbb{Z}) \wr S_r$ for all but finitely many $c \in \mathbb{Q}$.

Excluding this finite set we therefore know the structure of all the Galois groups $G_{n,c}$. The Chebotarev density theorem can then be used to determine the value of $\delta(T_{n,c})$ by a straightforward calculation within the group $(\mathbb{Z}/n\mathbb{Z}) \wr S_r$. In this way we obtain the following result.

Theorem 1.5. *There exists a finite set $E \subset \mathbb{Q}$ such that the following lower bounds hold for every $c \in \mathbb{Q} \setminus E$:*

$$\delta(T_{5,c}) > 0.81, \quad \delta(T_{6,c}) > 0.84, \quad \delta(T_{7,c}) > 0.86, \quad \delta(T_{9,c}) > 0.89.$$

The proof of Theorem 1.4 relies on Hilbert's irreducibility theorem and Faltings's theorem to reduce the proof to a problem of showing that certain algebraic curves have genera greater than 1. More precisely, let S be a splitting field of Φ_n over $\mathbb{Q}(t)$, so that $G_n = \mathrm{Gal}(S/\mathbb{Q}(t))$, and let $\mathcal{X}$ be the smooth projective curve over $\mathbb{Q}$ whose function field is S. As explained in Section 2, in order to show that the set E_n is finite it suffices to show that, for every maximal proper subgroup $M < G_n$, the quotient curve $\mathcal{X}/M$ has genus greater than 1. Our main objective is therefore to compute the genera of these quotient curves, or at least to obtain lower bounds for them.

The methods we develop for this purpose allow us to reduce the problem to a series of computations within the groups G_n. For $n \in \{5, 6\}$ we are able to determine the genera exactly, and for $n \in \{7, 9\}$ we prove lower bounds which suffice for our purposes. Though the methods used here could in principle be used to extend our results to higher values of n, there are computational limitations which prevent this. For instance, the group G_{11} has order $11^{186}(186)!$ and the cost of computing its maximal subgroups is prohibitively expensive. Other computational issues are discussed in Section 7.

Though it would be desirable to explicitly determine the finite sets E_n in Theorem 1.4, our method of proof does not suggest a feasible way of doing this. Indeed, one would have to determine the sets of rational points on several curves of very large genera, a problem which seems impossible with current methods. Nevertheless, in Section 9 we make some elementary observations regarding the sets E_n; for instance, they are always nonempty.

This article is organized as follows. In Section 2 we establish two foundational results for the rest of the article. In Section 3 we prove a theorem concerning the structure of inertia groups in Galois extensions of valued fields; this may be of independent interest. In Section 4 we recall various properties of dynatomic polynomials which were mostly proved by P. Morton. In Section 5 we study the action of G_n on the roots of Φ_n. In Sections 6 and 7 we apply the results of earlier sections to carry out the genus computations from which Theorem 1.4 can be deduced. In Section 8 we prove Theorem 1.5. Finally, in Section 9 we list the known elements of the sets E_n.

2. Preliminaries

Let n be a positive integer and let Φ_n be the polynomial defined in (1-1). Let S be a splitting field of Φ_n over $\mathbb{Q}(t)$, and $G_n = \mathrm{Gal}(S/\mathbb{Q}(t))$. Recall that E_n denotes the set of all rational numbers c such that $G_{n,c} \ncong G_n$, where $G_{n,c}$ is the Galois group of $\Phi_n(c, x)$ over $\mathbb{Q}$. The following lemma provides sufficient conditions for E_n to be a finite set.

Lemma 2.1. *Let $M_1, \dots, M_s$ be representatives of all the conjugacy classes of maximal subgroups of G_n, and let L_i denote the fixed field of M_i. Suppose that every function field L_i has genus greater than* 1. *Then E_n is finite.*

Proof. Let $\mathcal{X}$ be the smooth projective curve with function field S, and for every index i, let $\mathcal{X}_i$ be the quotient curve $\mathcal{X}/M_i$. It follows from the proof of Proposition 3.3.1 in [Serre 2008] (see also [Krumm and Sutherland 2017, Theorem 1.1]) that there exist a finite set $\mathcal{E} \subset \mathbb{P}^1(\mathbb{Q})$ and morphisms $\pi_i : \mathcal{X}_i \to \mathbb{P}^1$ such that

$$E_n \subseteq \mathcal{E} \cup \bigcup_{i=1}^{s} \pi_i(\mathcal{X}_i(\mathbb{Q})).$$

Since L_i is the function field of $\mathcal{X}_i$, the hypotheses imply that the smooth projective model of $\mathcal{X}_i$ has genus greater than 1, and hence, by Faltings's theorem [1983], the set $\mathcal{X}_i(\mathbb{Q})$ is finite. The result follows immediately. □

In view of Lemma 2.1, the main objects of interest in this article are the genera of the minimal intermediate fields in the extension $S/\mathbb{Q}(t)$. Our first step towards understanding these genera will be to show that in computing them we may replace $\mathbb{Q}$ with any subfield of $\mathbb{C}$.

Proposition 2.2. *Let $\mathbb{F}$ be any field satisfying $\mathbb{Q} \subseteq \mathbb{F} \subseteq \mathbb{C}$, and let N be a splitting field of Φ_n over $\mathbb{F}(t)$. Then there is an isomorphism*

$$\iota : \mathrm{Gal}(N/\mathbb{F}(t)) \longrightarrow \mathrm{Gal}(S/\mathbb{Q}(t))$$

with the following property: if A is a subgroup of $\mathrm{Gal}(N/\mathbb{F}(t))$ and $B = \iota(A)$, then the fixed fields of A and B have the same genus.

Proof. Let Σ be a splitting field of Φ_n over $\mathbb{C}(t)$, and let $R \subset \Sigma$ be the set of roots of Φ_n. By basic field theory, we may identify N with the field $\mathbb{F}(t)(R)$ and S with the field $\mathbb{Q}(t)(R)$. Restriction of automorphisms then yields injective homomorphisms

$$\mathrm{Gal}(\Sigma/\mathbb{C}(t)) \hookrightarrow \mathrm{Gal}(N/\mathbb{F}(t)) \hookrightarrow \mathrm{Gal}(S/\mathbb{Q}(t)). \tag{2-1}$$

The group $\mathrm{Gal}(\Sigma/\mathbb{C}(t))$ is naturally isomorphic to a subgroup $G_{\mathbb{C}}$ of the symmetric group $\mathrm{Sym}(R)$. (Explicitly, the isomorphism is given by restriction to R.) Similarly, we define groups $G_{\mathbb{F}}$ and $G_{\mathbb{Q}}$. By (2-1) we have

$$G_{\mathbb{C}} \leq G_{\mathbb{F}} \leq G_{\mathbb{Q}} \leq \mathrm{Sym}(R). \tag{2-2}$$

The polynomial $\phi(x) = x^2 + t$ permutes the elements of R (see, for instance, [Krumm 2016, §2.2]); thus we may regard ϕ as an element of the group $\mathrm{Sym}(R)$. Let $\mathcal{C}$ denote the centralizer of ϕ in $\mathrm{Sym}(R)$. Since ϕ is a polynomial map, it commutes with every element of $\mathrm{Gal}(S/\mathbb{Q}(t))$, and therefore $G_{\mathbb{Q}} \leq \mathcal{C}$. Now, by Theorem 3 in [Bousch 1992, Chapter 3] we have $G_{\mathbb{C}} = \mathcal{C}$. Hence, (2-2) implies that $G_{\mathbb{C}} = G_{\mathbb{F}} = G_{\mathbb{Q}}$. It follows that the embeddings (2-1) are in fact isomorphisms; in particular, restriction to S is an isomorphism

$$\iota : \mathrm{Gal}(N/\mathbb{F}(t)) \xrightarrow{\sim} \mathrm{Gal}(S/\mathbb{Q}(t)). \tag{2-3}$$

We now digress briefly from the main proof.

Lemma 2.3. *The field $\mathbb{Q}$ is algebraically closed in S.*

Proof. Let k be the algebraic closure of $\mathbb{Q}$ in S. By general theory of algebraic function fields, the extension $k/\mathbb{Q}$ is finite; moreover, it can easily be shown to be a Galois extension. To see that $k/\mathbb{Q}$ is normal, let $p(x) \in \mathbb{Q}[x]$ be an irreducible polynomial having a root in k. Then p remains irreducible in $\mathbb{Q}(t)[x]$ (see Lemma 3.1.10 in [Stichtenoth 2009]) and has a root in S; therefore p splits in S. However, by definition of k, every root of p in S belongs to k. Hence, p splits in k.

Since $k(t)$ is the composite of k and $\mathbb{Q}(t)$, the extension $k(t)/\mathbb{Q}(t)$ is Galois, and restriction to k yields an isomorphism

$$\mathrm{Gal}(k(t)/\mathbb{Q}(t)) \cong \mathrm{Gal}(k/k \cap \mathbb{Q}(t)) = \mathrm{Gal}(k/\mathbb{Q}).$$

It follows that there is a surjective homomorphism $\mathrm{Gal}(S/\mathbb{Q}(t)) \to \mathrm{Gal}(k/\mathbb{Q})$ with kernel $H := \mathrm{Gal}(S/k(t))$. Now, taking $\mathbb{F} = k$ in (2-3), the image of ι is clearly contained in H, so that in fact $H = \mathrm{Gal}(S/\mathbb{Q}(t))$. Therefore $\mathrm{Gal}(k/\mathbb{Q})$ must be trivial, and $k = \mathbb{Q}$. □

Returning to the proof of the proposition, let $A \leq \mathrm{Gal}(N/\mathbb{F}(t))$ and set $B = \iota(A)$. Let U and V be the fixed fields of A and B, respectively. Thus, U and V are intermediate fields in the extensions $N/\mathbb{F}(t)$ and $S/\mathbb{Q}(t)$. We claim that U is the composite of V and $\mathbb{F}$. The fact that $U \supseteq V$ follows immediately from the definitions, and it is clear that $U \supseteq \mathbb{F}$; hence $U \supseteq V\mathbb{F}$. To prove that $U = V\mathbb{F}$ we will show that $[U : \mathbb{F}(t)] = [V\mathbb{F} : \mathbb{F}(t)]$. Since ι is an isomorphism mapping A to B, we have

$$[U : \mathbb{F}(t)] = |\mathrm{Gal}(N/\mathbb{F}(t)) : A| = |\mathrm{Gal}(S/\mathbb{Q}(t)) : B| = [V : \mathbb{Q}(t)].$$

Thus, it suffices to show that $[V : \mathbb{Q}(t)] = [V\mathbb{F} : \mathbb{F}(t)]$. Let α be a primitive element for V over $\mathbb{Q}(t)$, and let $p \in \mathbb{Q}(t)[x]$ be the minimal polynomial of α. Clearly $V\mathbb{F} = \mathbb{F}(t)(\alpha)$, so it is enough to show that p remains irreducible over $\mathbb{F}(t)$. Since p is irreducible over $\mathbb{Q}(t)$, the group $\mathrm{Gal}(S/\mathbb{Q}(t))$ acts transitively on the roots of p. This, together with the fact that ι is given by restriction to N, imply that $\mathrm{Gal}(N/\mathbb{F}(t))$ also acts transitively on the roots of p, and therefore p is irreducible over $\mathbb{F}(t)$. This completes the proof that $U = V\mathbb{F}$.

It remains only to show that U and V have the same genus. Since $\mathbb{F}$ contains the constant field of V (by Lemma 2.3), $U = V\mathbb{F}$ is a constant field extension of V (in the terminology of [Stichtenoth 2009, §3.6]). Equality between the genera of U and V now follows from Theorem 22 in [Artin 2006, p. 291]; see also Theorem 3.6.3 in [Stichtenoth 2009]. □

From Lemma 2.1 and Proposition 2.2 we deduce the following proposition, which is the key result of this section.

Proposition 2.4. *Let N be a splitting field of Φ_n over $\overline{\mathbb{Q}}(t)$. Let $M_1, \ldots, M_s$ be representatives of all the conjugacy classes of maximal subgroups of the group $G = \mathrm{Gal}(N/\overline{\mathbb{Q}}(t))$, and let L_i be the fixed field of M_i. Suppose that the genus of L_i is greater than 1 for every index i. Then the set E_n is finite.*

3. A result in valuation theory

Let K be a field, and let $v : K^* \to \mathbb{R}$ be a discrete valuation of K with perfect residue field k. Let N be a finite Galois extension of K with Galois group $G = \mathrm{Gal}(N/K)$. For any elements $\sigma, \tau \in G$ we will write τ^σ to denote the conjugate $\sigma^{-1}\tau\sigma$; similarly, for any subgroup $A \leq G$ we let $A^\sigma = \sigma^{-1}A\sigma$.

If L is an intermediate field in the extension N/K and w is a valuation of N extending a valuation u of L, we denote by $D_{w|u}$ and $I_{w|u}$ the decomposition and inertia groups of w over u. If u extends the valuation v of K, we let $e_{u|v}$ and $f_{u|v}$ denote the ramification index and residue degree of u over v.

Lemma 3.1. *Let w be a valuation of N extending v, and let $D = D_{w|v}$ and $I = I_{w|v}$. Let H be a subgroup of G with fixed field L, and let S_L be the set of all valuations of L extending v. Then there is a well-defined bijection*

$$D\backslash G/H \xrightarrow{\sim} S_L$$

given by $D\sigma H \mapsto (w \circ \sigma)|_L$. Furthermore, if $u = (w \circ \sigma)|_L$, then

$$e_{u|v} \cdot f_{u|v} = |D^\sigma : D^\sigma \cap H| \quad \textit{and} \quad e_{u|v} = |I^\sigma : I^\sigma \cap H|. \tag{3-1}$$

Proof. The first statement is well known; a proof may be found in Lemma 17.1.2 and Corollary 17.1.3 of [Efrat 2006]. Suppose now that $u = (w \circ \sigma)|_L$, and let $\tilde{w} = w \circ \sigma$. It is then a simple exercise to show that

$$D_{\tilde{w}|u} = D^\sigma \cap H \quad \text{and} \quad I_{\tilde{w}|u} = I^\sigma \cap H. \tag{3-2}$$

Note that $D^\sigma = D_{\tilde{w}|v}$ and $I^\sigma = I_{\tilde{w}|v}$. Now, since k is perfect, we have $|D_{\tilde{w}|v}| = e_{\tilde{w}|v} \cdot f_{\tilde{w}|v}$ and $|I_{\tilde{w}|v}| = e_{\tilde{w}|v}$ (see [Neukirch 1999, Chapter I, Proposition 9.6]). The relations (3-1) now follow easily from (3-2). □

Proposition 3.2. *Suppose that N is the splitting field of an irreducible polynomial $P(x) \in K[x]$. Let F be a subextension of N/K obtained by adjoining one root of $P(x)$ to K. Let $u_1, \ldots, u_m$ be the distinct valuations of F extending v, and set $e_i = e_{u_i|v}$ and $f_i = f_{u_i|v}$. Let w be a valuation of N extending v, and assume that $e_{w|v}$ is not divisible by the characteristic of k. Then the inertia group $I_{w|v}$ is generated by an element whose disjoint cycle decomposition (as a permutation of the roots of P) has the form*

$$\underbrace{(e_1\text{-cycle}) \cdots (e_1\text{-cycle})}_{f_1 \text{ times}} \cdots \underbrace{(e_m\text{-cycle}) \cdots (e_m\text{-cycle})}_{f_m \text{ times}}. \tag{3-3}$$

Proof. Set $D = D_{w|v}$ and $I = I_{w|v}$. The assumption that the characteristic of k does not divide $|I|$ implies that I is a cyclic group; see [Stichtenoth 2009, Proposition 3.8.5] or [Efrat 2006, §16.2]. Let R denote the set of roots of $P(x)$ in N, and consider the natural action of I on R. Let $\mathcal{O}$ be the set of orbits of this action. We will show that $\mathcal{O}$ can be partitioned into subsets $S_1, \ldots, S_m$ such that every orbit in S_i has cardinality e_i, and $\#S_i = f_i$. Note that this implies that every generator of I has a cycle decomposition of the form (3-3).

For every $x \in R$ let $\mathcal{O}_x$ and I_x, respectively, denote the orbit of x (under the action of I) and the stabilizer of x in I. Let $r \in R$ be such that $F = K(r)$, and set $H = \mathrm{Gal}(N/F)$. Note that H is the stabilizer of r in G.

By Lemma 3.1, there exist distinct double cosets $D\sigma_1 H, \dots, D\sigma_m H$ such that $u_i = (w \circ \sigma_i)|_F$. For $i = 1, \dots, m$ we define a map ψ_i as follows:

$$I^{\sigma_i} \backslash D^{\sigma_i} / (D^{\sigma_i} \cap H) \xrightarrow{\psi_i} \mathcal{O},$$
$$I^{\sigma_i} \tau (D^{\sigma_i} \cap H) \longmapsto \mathcal{O}_{\sigma_i \tau(r)}.$$

A straightforward calculation shows that ψ_i is well defined and injective. Letting $S_i \subseteq \mathcal{O}$ be the image of ψ_i, we claim that the sets $S_1, \dots, S_m$ have the properties stated above.

We begin by showing that every orbit in S_i has cardinality e_i. To ease notation, let us fix an index i and set $\sigma = \sigma_i$ and $M = D^\sigma \cap H$. Letting $\tau \in D^\sigma$, we must show that $\#\mathcal{O}_{\sigma\tau(r)} = e_i$. Note that $(I_{\sigma\tau(r)})^{\sigma\tau} = I^{\sigma\tau} \cap H$, so that $|I_{\sigma\tau(r)}| = |I^{\sigma\tau} \cap H|$, and therefore

$$\#\mathcal{O}_{\sigma\tau(r)} = |I : I_{\sigma\tau(r)}| = \frac{|I|}{|I_{\sigma\tau(r)}|} = \frac{|I^{\sigma\tau}|}{|I_{\sigma\tau(r)}|} = \frac{|I^{\sigma\tau}|}{|I_{\sigma\tau} \cap H|} = |I^{\sigma\tau} : I^{\sigma\tau} \cap H|.$$

Now, since $\tau \in D^\sigma$, we have $\sigma\tau \in D\sigma H$. Lemma 3.1 then implies that $(w \circ \sigma\tau)|_F = (w \circ \sigma)|_F = u_i$ and $|I^{\sigma\tau} : I^{\sigma\tau} \cap H| = e_i$. Hence $\#\mathcal{O}_{\sigma\tau(r)} = e_i$.

Next we show that $\#S_i = f_i$. Note that $\#S_i = \#I^\sigma \backslash D^\sigma / M$ since ψ_i is injective. The fact that I is a normal subgroup of D implies that

$$I^\sigma \backslash D^\sigma / M = D^\sigma / (I^\sigma M).$$

Thus, using Lemma 3.1 we obtain

$$\#S_i = |D^\sigma| / |I^\sigma M| = \frac{|D^\sigma| \cdot |I^\sigma \cap H|}{|D^\sigma \cap H| \cdot |I^\sigma|} = \frac{|D^\sigma : D^\sigma \cap H|}{|I^\sigma : I^\sigma \cap H|} = \frac{e_i f_i}{e_i} = f_i.$$

Now we show that the sets $S_1, \dots, S_m$ are pairwise disjoint. Suppose, by contradiction, that there exist distinct indices i, j such that $S_i \cap S_j \neq \varnothing$. Then there exist $\alpha \in D^{\sigma_i}$, $\beta \in D^{\sigma_j}$, and $\gamma \in I$ such that $\sigma_i \alpha(r) = \gamma \sigma_j \beta(r)$. Writing $\alpha = \sigma_i^{-1} \delta \sigma_i$ and $\beta = \sigma_j^{-1} d \sigma_j$ with $\delta, d \in D$, this implies that $\delta\sigma_i(r) = \gamma d \sigma_j(r)$; hence, there exists $h \in H$ such that $\sigma_i = \delta^{-1}\gamma d \sigma_j h$. Note that $\delta^{-1}\gamma d \in D$, so the previous equality implies that $\sigma_i \in D\sigma_j H$ and therefore $D\sigma_i H = D\sigma_j H$, a contradiction.

Finally, we show that $\mathcal{O} = \bigcup_{i=1}^m S_i$. Let $R_1, \dots, R_m$ be the subsets of R defined by $R_i = \bigcup_{C \in S_i} C$. From the results proved above it follows that $\#R_i = e_i f_i$ and that the sets $R_1, \dots, R_m$ are pairwise disjoint. Given that v is a discrete valuation, we have the relation $[F : K] = \sum_{i=1}^m e_i f_i$. Hence

$$\#R = \deg(P) = [F : K] = \sum_{i=1}^m e_i f_i = \sum_{i=1}^m \#R_i = \#\bigcup_{i=1}^m R_i.$$

It follows that $R = \bigcup_{i=1}^m R_i$, which implies that $\mathcal{O} = \bigcup_{i=1}^m S_i$. □

Remark 3.3. Proposition 3.2 was inspired by a theorem of Beckmann [1994] concerning inertia groups in Galois extensions of $\mathbb{Q}$; indeed, Beckmann's result is essentially the case $K = \mathbb{Q}$ of the proposition. However, the proof given here has little in common with the proof in [loc. cit.].

Proposition 3.4. *With notation and assumptions as in Proposition 3.2, let γ be a generator of $I_{w|v}$ and let H be a subgroup of G with fixed field L. Suppose that $D_{w|v} = I_{w|v}$. Then the number of valuations u of L extending v such that $e_{u|v} = 1$ is given by*

$$\frac{|C_G(\gamma)| \cdot s(H, \gamma)}{|H|}, \tag{3-4}$$

where $C_G(\gamma)$ is the centralizer of γ in G and $s(H, \gamma)$ is the number of G-conjugates of γ that belong to H.

Proof. Let $D = D_{w|v}$ and define sets $A = \{\sigma \in G \mid \gamma^\sigma \in H\}$ and

$$\Delta = \{D\sigma H \in D\backslash G/H \mid \sigma \in A\}.$$

It follows from Lemma 3.1 that the cardinality of Δ is equal to the number of valuations u of L extending v such that $e_{u|v} = 1$. Thus, in order to prove the proposition it suffices to show that $|H| \cdot (\#\Delta) = |C_G(\gamma)| \cdot s(H, \gamma)$.

For every element $a \in A$ the right coset $C_G(\gamma)\cdot a$ is contained in A; hence, the set $U = \{C_G(\gamma)\cdot a \mid a \in A\}$ is a partition of A into subsets of size $|C_G(\gamma)|$. Thus $\#A = |C_G(\gamma)| \cdot (\#U)$. Now let $B = \{\gamma^\sigma \mid \sigma \in G\} \cap H$, so that $\#B = s(H, \gamma)$. Note that $\#U = \#B$; indeed, there is a bijective map $U \to B$ given by $C_G(\gamma)\cdot a \mapsto \gamma^a$. Therefore,

$$\#A = |C_G(\gamma)| \cdot (\#B) = |C_G(\gamma)| \cdot s(H, \gamma). \tag{3-5}$$

Let $f : A \twoheadrightarrow \Delta$ be the surjective map given by $f(\sigma) = D\sigma H$. We claim that, for every $a \in A$, $f^{-1}(f(a)) = aH$. It is clear that $aH \subseteq f^{-1}(f(a))$. Now suppose that $f(a') = f(a)$, so that $a' = dah$ for some $d \in D$ and $h \in H$. Since $\gamma^a \in H$, we may write $\gamma a = ah'$ for some $h' \in H$. Furthermore, since $D = I_{w|v} = \langle\gamma\rangle$, we have $d = \gamma^n$ for some positive integer n. Thus

$$a' = dah = \gamma^n ah = a(h')^n h \in aH,$$

which proves the claim. Since every fiber of f has cardinality $|H|$, we have $\#A = |H| \cdot (\#\Delta)$, and hence, by (3-5), $|H| \cdot (\#\Delta) = |C_G(\gamma)| \cdot s(H, \gamma)$. □

For later reference, we include here a combined statement of Propositions 3.2 and 3.4 in the special case where K is the function field $\overline{\mathbb{Q}}(t)$ and the valuation v corresponds to a place p of K. Note that in this case all residue degrees $f_{u|v}$ are equal to 1.

Corollary 3.5. *Let t be an indeterminate and $K = \overline{\mathbb{Q}}(t)$. Suppose that $P(x) \in K[x]$ is irreducible, and let N be a splitting field for $P(x)$. Let F be a subextension of N/K obtained by adjoining one root of $P(x)$ to K. Let p be a place of K, and let $\mathfrak{p}_1, \ldots, \mathfrak{p}_m$ be the distinct places of F lying over p. Then, for every place $\mathfrak{P}$ of N lying over p, the inertia group $I_{\mathfrak{P}|p}$ is generated by an element γ whose disjoint cycle decomposition has the form $(e_1\text{-cycle}) \cdots (e_m\text{-cycle})$, where e_i is the ramification index of $\mathfrak{p}_i$ over p. Furthermore, if H is a subgroup of $G = \mathrm{Gal}(N/K)$ with fixed field L, then the number of places of L lying over p which are unramified over K is given by the formula (3-4).*

4. Ramification data for dynatomic polynomials

Let us fix a positive integer n. We will henceforth regard the polynomial $\Phi_n(x)$ as an element of the ring $\overline{\mathbb{Q}}(t)[x]$. As such, it is known by work of Bousch [1992, Chapter 3] that Φ_n is irreducible. In this section we will apply Corollary 3.5 to study inertia groups in the Galois group of Φ_n.

Let $K = \overline{\mathbb{Q}}(t)$, let N/K be a splitting field of Φ_n, and let $G = \mathrm{Gal}(N/K)$. Let F be a subextension of N/K obtained by adjoining one root of Φ_n to K. Morton [1996, §3] studies the ramification of places in the extension F/K by using certain polynomials $\Delta_{n,d} \in \mathbb{Z}[t]$, where d is a divisor of n. These polynomials had previously been defined in [Morton and Vivaldi 1995, §1]; we refer the reader to that article for the definition. We now recall a few results from [Morton 1996; Morton and Vivaldi 1995] which will be needed here.

For every positive integer s, let

$$\nu(s) = \tfrac{1}{2} \sum_{d \mid s} \mu(s/d) 2^d.$$

Lemma 4.1 (Morton–Vivaldi). *For every divisor d of n, let $R_{n,d} \subset \overline{\mathbb{Q}}$ denote the set of roots of $\Delta_{n,d}$. Then the following hold*:

(a) $\#R_{n,d} = \deg \Delta_{n,d}$ *for every* d.

(b) *If d and e are distinct divisors of n, then $R_{n,d} \cap R_{n,e} = \varnothing$.*

(c) *Letting φ denote Euler's phi function, the degree of $\Delta_{n,d}$ is given by*

$$\deg \Delta_{n,d} = \begin{cases} \nu(d)\varphi(n/d) & \text{if } d < n, \\ \nu(n) - \sum_{\substack{k \mid n \\ k<n}} \nu(k)\varphi(n/k) & \text{if } d = n. \end{cases}$$

Proof. All statements are proved in [Morton and Vivaldi 1995]. Indeed, (a) and (b) follow from Proposition 3.2, and (c) follows from Corollary 3.3. □

Recall that for every place p of K, the *conorm* of p with respect to the extension F/K is the divisor, which we write multiplicatively, defined by

$$i_{F/K}(p) = \mathfrak{p}_1^{e_1} \cdots \mathfrak{p}_s^{e_s},$$

where $\mathfrak{p}_1, \ldots, \mathfrak{p}_s$ are the distinct places of F lying over p and e_i is the ramification index of $\mathfrak{p}_i$ over p. A discussion of the basic properties of the conorm map may be found in [Stichtenoth 2009, §3.1] or [Rosen 2002, Chapter 7].

Let $D = \deg \Phi_n$; note that $D = 2\nu(n)$. As explained in Section 5, the set of roots of Φ_n can be partitioned into sets of cardinality n, and therefore n divides D. Let $r = D/n$.

Lemma 4.2 (Morton). *Let p_∞ be the infinite place of K, i.e., the place corresponding to the valuation v_∞ of K given by $v_\infty(f/g) = \deg g - \deg f$. For $b \in \overline{\mathbb{Q}}$, let p_b denote the place of K corresponding to the polynomial $t - b$.*

(a) *The places of K that ramify in F are p_∞ and p_b for $b \in \bigcup_{d \mid n} R_{n,d}$.*

(b) *The conorm of p_∞ has the form*

$$i_{F/K}(p_\infty) = \mathfrak{p}_1^2 \cdots \mathfrak{p}_{\nu(n)}^2.$$

(c) *For every $b \in R_{n,n}$, the conorm of p_b has the form*

$$i_{F/K}(p_b) = \mathfrak{p}_1^2 \cdots \mathfrak{p}_n^2 \cdot \mathfrak{q}_1 \cdots \mathfrak{q}_{n(r-2)}.$$

(d) *For every $b \in R_{n,d}$, where $d < n$, the conorm of p_b has the form*

$$i_{F/K}(p_b) = \mathfrak{p}_1^{n/d} \cdots \mathfrak{p}_d^{n/d} \cdot \mathfrak{q}_1 \cdots \mathfrak{q}_{n(r-1)}.$$

Proof. All statements are proved in [Morton 1996]; (a), (c) and (d) follow from the proof of Proposition 9, and (b) follows from Proposition 10. □

Let $\mathbb{P} = \{p_\infty\} \cup \left\{p_b \mid b \in \bigcup_{d \mid n} R_{n,d}\right\}$ be the set of places of K that ramify in F. For any intermediate field L in the extension N/K and any place p of K, let $\mathbb{P}_L(p)$ denote the set of places of L lying over p.

We introduce some terminology to be used throughout the article. Suppose that G is a group acting on a finite set X, and let $g \in G$. We say that g has *cycle type* (a, b), where a and b are positive integers, if the disjoint cycle decomposition of g, disregarding 1-cycles, is a product of b a-cycles.

Applying Corollary 3.5 to the polynomial Φ_n and using Lemma 4.2, we immediately obtain the following description of inertia groups in G.

Proposition 4.3. *Let $p \in \mathbb{P}$ and $\mathfrak{P} \in \mathbb{P}_N(p)$. Then the inertia group $I_{\mathfrak{P} \mid p}$ has a generator with cycle type (a, b) satisfying*

$$(a, b) = \begin{cases} (2, D/2) & \text{if } p = p_\infty, \\ (2, n) & \text{if } p = p_b \text{ with } b \in R_{n,n}, \\ (n/d, d) & \text{if } p = p_b \text{ with } b \in R_{n,d},\ d < n. \end{cases}$$

In addition to the data on ramification of places in F/K provided by Lemma 4.2, in later sections we will need some ramification data for a subfield $F_0 \subset F$ defined as follows. Let θ be a root of Φ_n such that $F = K(\theta)$. The field F has an automorphism[1] given by $\theta \mapsto \phi(\theta) = \theta^2 + t$; we define F_0 to be the fixed field of this automorphism.

Proposition 4.4 (Morton). *Let $p \in \mathbb{P}$ and let $S(p) = \sum_{\mathfrak{q} \in \mathbb{P}_{F_0}(p)} (e_{\mathfrak{q} \mid p} - 1)$.*

(a) *If $p = p_\infty$, then $S(p) = r - e_n$, where*

$$e_n = \frac{1}{2n} \sum_{d \mid (n,2)} \varphi(d)^2 \cdot \sum_{k \in U_{n,d}} \mu(n/k) 2^{k/d}.$$

Here $U_{n,d} = \{k \in \mathbb{Z}_{>0} : k \mid n, d \mid k, \text{ and } (n/k, d) = 1\}$.

(b) *If $p = p_b$, where $b \in R_{n,n}$, then $S(p) = 1$.*

(c) *If $p = p_b$, where $b \in R_{n,d}$ for some $d < n$, then $S(p) = 0$.*

[1] Note that $\phi(\theta)$ is a root of Φ_n, so there is an isomorphism $F \to K(\phi(\theta))$ mapping θ to $\phi(\theta)$. Moreover, the fact that $\phi^n(\theta) = \theta$ implies that $F = K(\phi(\theta))$, so this map is in fact an automorphism of F.

Proof. All statements are proved in [Morton 1996]; (a) follows Theorem 13, while (b) and (c) can be deduced from the proof of Proposition 9. Indeed, it is shown in that proposition that if $p = p_b$, where $b \in R_{n,n}$, then there is a unique ramified place of F_0 lying over p, and its ramification index is 2; this implies (b). Similarly, if $p = p_b$, where $b \in R_{n,d}$ for some $d < n$, then p is unramified in F_0, which implies (c). □

5. The action of the Galois group of Φ_n

We continue using here the notation introduced in the previous section. The genus computations in Sections 6 and 7, which form the core of this article, rely fundamentally on Propositions 3.4 and 4.3. In order to apply these propositions effectively, we require a precise understanding of the elements of G whose cycle decompositions have the forms described in Proposition 4.3. In addition, explicit formulas for the orders of the centralizers of these elements will be needed when applying Proposition 3.4. The purpose of this section is to carry out a detailed analysis of the action of G on the roots of Φ_n. In the process we address both of the above requirements, the key result being Proposition 5.5.

Recall the notion of an isomorphism of group actions: if A and B are groups acting on sets X and Y, respectively, we write $A \equiv B$ if there exist a group isomorphism $\varphi : A \to B$ and a bijection $\varepsilon : X \to Y$ such that $\varepsilon(ax) = \varphi(a)\varepsilon(x)$ for all $a \in A$ and $x \in X$. Though the notation $A \equiv B$ does not make reference to the sets X and Y, this should cause no confusion here because the sets being acted on will be clear from context.

Let R be the set of roots of Φ_n in the splitting field N, and consider the natural action of G on R. In this section we will discuss three group actions, which we refer to as *realizations* of G, that are isomorphic to G with its action on R. The first realization is the automorphism group of a graph acting on its set of vertices; this is helpful as a visual aid for understanding the action of G. The second realization is a particular subgroup of the symmetric group S_D acting on the set $\{1, \dots, D\}$; this is useful for carrying out explicit computations with elements of G. The third realization is a wreath product $(\mathbb{Z}/n\mathbb{Z}) \wr S_r$ acting on the set $(\mathbb{Z}/n\mathbb{Z}) \times \{1, \dots, r\}$. Though somewhat more technical, we find that this realization is the most convenient for purposes of proving the main results of this section. The key fact needed to show that these realizations are isomorphic is a well-known theorem of Bousch [1992, Chapter 3], namely Theorem 3.

5A. ***The group G as a graph automorphism group.*** It is a simple consequence of the definition of Φ_n that the map $\phi(x) = x^2 + t$ permutes the elements of R (see [Krumm 2016, §2.2] for details). Regarding ϕ as an element of the symmetric group $\mathrm{Sym}(R)$, we may therefore partition the set R into ϕ-orbits. By [Morton and Patel 1994, Theorem 2.4(c)], the fact that Φ_n is irreducible implies that every orbit has size n; hence, the number of orbits is $(\#R)/n = D/n = r$.

Let $\mathcal{G}$ be the natural embedding of G in $\mathrm{Sym}(R)$, and note that $G \equiv \mathcal{G}$. Let Γ be the directed graph whose vertices are the elements of R and which has an edge $x \to \phi(x)$ for every $x \in R$. An illustration of Γ is shown in Figure 1 below. By Bousch's theorem, $\mathcal{G}$ is the centralizer of ϕ in $\mathrm{Sym}(R)$. (More explicitly, this is a consequence of the proof of Proposition 2.2. In the notation of that proof, we have $\mathcal{G} = G_{\mathbb{F}}$, where $\mathbb{F} = \overline{\mathbb{Q}}$.) It follows that $\mathcal{G} = \mathrm{Aut}(\Gamma)$ and therefore $G \equiv \mathrm{Aut}(\Gamma)$.

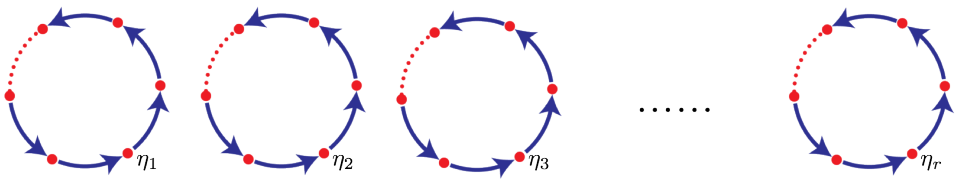


Figure 1. A directed graph whose automorphism group is isomorphic to the Galois group of Φ_n. Every cycle in the graph has n vertices, and there are r cycles in total.

5B. ***The group G as a permutation group.*** Let S_D be the symmetric group on the set $\{1, \dots, D\}$ and let $\sigma \in S_D$ be the permutation defined by

$$\sigma = (1, \dots, n)(n+1, \dots, 2n) \cdots (D-n+1, \dots, D).$$

There is a bijection $\ell : \{1, \dots, D\} \to R$ under which the cycles in the decomposition of σ correspond to the cycles in the graph Γ. Indeed, if we choose representatives $\eta_1, \dots, \eta_r$ of the distinct cycles in Γ, then one such map ℓ is given by

$$\ell(ni - j) = \phi^{n-j}(\eta_i) \quad \text{for } 1 \le i \le r \text{ and } 0 \le j < n.$$

The map ℓ induces an isomorphism $\iota : S_D \to \mathrm{Sym}(R)$ under which σ maps to ϕ. Let $\mathcal{Z}$ be the centralizer of σ in S_D. Since $\mathcal{G}$ is the centralizer of ϕ in $\mathrm{Sym}(R)$, the image of $\mathcal{Z}$ under ι is equal to $\mathcal{G}$. Moreover, the maps ι and ℓ induce an isomorphism of group actions between $\mathcal{Z}$ and $\mathcal{G}$; hence $G \equiv \mathcal{Z}$.

5C. ***Background on wreath products.*** Before discussing the realization of G as a wreath product, we recall the basic construction of wreath products. For further information on this topic we refer the reader to [Dixon and Mortimer 1996, §2.6; Rotman 1995, Chapter 7; Kerber 1971, Chapter I].

Let S_r denote the symmetric group on the set $\Omega = \{1, \dots, r\}$. Let A be a group, and consider the direct product A^r consisting of functions $f : \Omega \to A$ with pointwise multiplication. There is an action of S_r on A^r given by $\pi \cdot f = f_\pi$, where f_π is the function

$$f_\pi(i) = f(\pi^{-1}(i)) \quad \text{for every } i \in \Omega.$$

This action induces a homomorphism $S_r \to \mathrm{Aut}(A^r)$, so we may form the semidirect product $\mathcal{W} = A^r \rtimes S_r$. Elements of $\mathcal{W}$ have the form (f, π), where $f \in A^r$ and $\pi \in S_r$; the group operation in $\mathcal{W}$ is given by

$$(f, \pi)(g, \sigma) = (f g_\pi, \pi\sigma).$$

The group $\mathcal{W}$ is the wreath product of A with S_r, denoted $A \wr S_r$. Letting e and 1, respectively, denote the identity elements of A^r and S_r, there are embeddings $A^r \hookrightarrow \mathcal{W}$ and $S_r \hookrightarrow \mathcal{W}$ given by $f \mapsto (f, 1)$ and $\pi \mapsto (e, \pi)$; we will henceforth identify A^r and S_r with their images under these maps. The group $B = A^r$, called the *base group* of the wreath product, is a normal subgroup of $\mathcal{W}$; indeed, B is the kernel

of the projection map $\mathcal{W} \to S_r$ given by $(f, \pi) \mapsto \pi$. Furthermore, S_r is a complement for B in the sense that $B \cap S_r$ is trivial and $BS_r = \mathcal{W}$.

Suppose now that A acts on a set Δ. Then there is an action of $\mathcal{W}$ on the Cartesian product $\Delta \times \Omega$ given by

$$(f, \pi) \cdot (d, i) = (f(\pi(i)) \cdot d, \pi(i)). \tag{5-1}$$

Moreover, this action is faithful if A acts faithfully on Δ.

5D. ***The group G as a wreath product.*** For the remainder of this section we assume that $A = \mathbb{Z}/n\mathbb{Z}$, so that $\mathcal{W} = (\mathbb{Z}/n\mathbb{Z}) \wr S_r$. The action of A on itself by addition induces a faithful action of $\mathcal{W}$ on the set $X = A \times \Omega$ given by (5-1). We will show that $\mathcal{W} \equiv G$.

Let $\eta_1, \ldots, \eta_r$ be representatives of the distinct ϕ-orbits of R. For every $w = (f, \pi) \in \mathcal{W}$ we define $\zeta_w \in \mathcal{G} = \mathrm{Aut}(\Gamma)$ by

$$\zeta_w(\phi^a(\eta_i)) = \phi^{f(\pi(i))+a}(\eta_{\pi(i)}) \quad \text{for } a \in A \text{ and } i \in \Omega.$$

Note that the notation ϕ^a for $a \in A$ is unambiguous since ϕ^n is the identity element of $\mathrm{Sym}(R)$. Using the fact that $\mathcal{G}$ is the centralizer of ϕ in $\mathrm{Sym}(R)$, it is a simple exercise to show that ζ_w is a well-defined element of $\mathcal{G}$, and that the map $\zeta : \mathcal{W} \to \mathcal{G}$ given by $w \mapsto \zeta_w$ is a group isomorphism.

Let $\varepsilon : X \to R$ be the map defined by $\varepsilon(a, i) = \phi^a(\eta_i)$. From the definitions it follows that ε is a bijection and that for every $w \in \mathcal{W}$ and $\alpha \in X$ we have $\varepsilon(w\alpha) = \zeta(w)\varepsilon(\alpha)$. Hence $\mathcal{W} \equiv \mathcal{G}$, and therefore $G \equiv \mathcal{W}$. Using this realization of G as a wreath product, we will now study the action of G.

Remark 5.1. It follows from the above discussion that

$$\mathrm{Aut}(\Gamma) \cong (\mathbb{Z}/n\mathbb{Z}) \wr S_r.$$

This is a special case of a well-known theorem of Frucht in graph theory. As shown in [Frucht 1949] (see also [Harary 1969, Theorem 14.5]), if Λ is a finite connected graph and Γ is a graph consisting of r disjoint copies of Λ, then $\mathrm{Aut}(\Gamma) \cong \mathrm{Aut}(\Lambda) \wr S_r$.

5E. ***Conjugacy in $\mathcal{W}$.*** Our main reason for using the realization of G as a wreath product is that it provides convenient ways of deciding whether two elements of G are conjugates of each other, and of calculating the order of the centralizer of any element of G. The key notion needed for these tasks is the *type* of an element of $\mathcal{W}$, defined below.

For every cycle $C = (i_1, i_2, \ldots, i_k) \in S_r$ and every element $f \in A^r$, we denote by $f(C)$ the element of A given by $f(C) = f(i_1) + \cdots + f(i_k)$.

For every element $w = (f, \pi) \in \mathcal{W}$, we define a map $T_w : X \to \mathbb{Z}_{\geq 0}$ as follows: for $a \in A$ and $k \in \Omega$, $T_w(a, k)$ is the number of k-cycles C in the cycle decomposition of π such that $f(C) = a$. The map T_w will be called the *type* of w. When w is clear from context, we will denote $T_w(a, k)$ simply by t_{ak} and we will use matrix notation (t_{ak}) to denote the map T_w.

Proposition 5.2. (1) *Let $w_1, w_2 \in \mathcal{W}$. Then w_1 and w_2 are conjugates if and only if they have the same type.*

(2) *If w has type (t_{ak}), then the order of the centralizer of w in $\mathcal{W}$ is given by the formula*

$$\prod_{a\in A}\prod_{k\in\Omega}(t_{ak})!(kn)^{t_{ak}}.$$

Proof. Both statements can be deduced from more general results proved in [Kerber 1971]. Specifically, (1) follows from item 3.7 on page 44, and (2) follows from item 3.9 on page 47. □

5F. ***The action of $\mathcal{W}$.*** In this section we prove various properties of the action of $\mathcal{W}$ on X. For elements $w = (f, \pi) \in \mathcal{W}$ and $\alpha = (a, i) \in X$, we will denote by $w(\alpha)$ the action of w on α. Thus,

$$w(\alpha) = (f(\pi(i)) + a, \pi(i)). \tag{5-2}$$

Let $C_i = A \times \{i\}$ for $1 \le i \le r$. Under the map ε defined in Section 5D, C_i corresponds to the i-th cycle in the graph Γ, i.e., the cycle containing η_i.

The base group $A^r \le \mathcal{W}$ is generated by the elements $\rho_1, \dots, \rho_r$ defined by $\rho_i = (\delta_i, 1)$, where $\delta_i(j) = 0$ if $j \ne i$ and $\delta_i(i) = 1$. Note that ρ_i maps C_i to itself and acts as the identity on C_j if $j \ne i$. Viewed as an element of $\mathrm{Aut}(\Gamma)$ (via the map ζ defined in Section 5D), ρ_i acts as a $1/n$ rotation on the i-th cycle. Let $\rho = \rho_1 \cdots \rho_r = (\delta, 1)$, where $\delta(i) = 1$ for all $i \in \Omega$. Then $\zeta(\rho) = \phi$, so ρ is in the center of $\mathcal{W}$. A simple calculation shows that for all $s \in \mathbb{Z}$, $a \in A$, and $i \in \Omega$ we have

$$\rho^s(a, i) = \rho_i^s(a, i) = (a + \bar{s}, i). \tag{5-3}$$

For every $w \in \mathcal{W}$ and every $i \in \Omega$, let $w(C_i) = \{w(\alpha) \mid \alpha \in C_i\}$.

Lemma 5.3. *Let $w = (f, \pi) \in \mathcal{W}$ and let $i \in \Omega$.*

(1) *Letting $j = \pi(i)$, we have $w(C_i) = C_j$.*

(2) *If $w(C_i) = C_i$, then there exists $0 \le s < n$ such that $w(\alpha) = \rho_i^s(\alpha)$ for every $\alpha \in C_i$. Moreover, the w-orbit of every element of C_i has cardinality $n/\gcd(n, s)$.*

Proof. For every element $(a, i) \in C_i$ we have $w(a, i) = (f(j) + a, j) \in C_j$, so $w(C_i) \subseteq C_j$. Since $\#C_i = \#C_j$ and w acts as a bijection on X, this implies that $w(C_i) = C_j$, proving (1). Suppose now that $w(C_i) = C_i$, and let $0 \le s < n$ be such that $w(0, i) = (\bar{s}, i)$. By (5-3) we have $w(0, i) = \rho^s(0, i)$. Given $\alpha \in C_i$, we may write α in the form $\alpha = (\bar{k}, i) = \rho^k(0, i)$ for some integer k. Using (5-3) and the fact that w commutes with ρ we obtain

$$w(\alpha) = w\rho^k(0, i) = \rho^k w(0, i) = \rho^k\rho^s(0, i) = \rho^s\rho^k(0, i) = \rho^s(\alpha) = \rho_i^s(\alpha).$$

This proves the first statement in (2). Since w acts like ρ_i^s on C_i, the orbit of α under w is equal to its orbit under ρ_i^s. The cyclic group generated by ρ_i^s has order $n/\gcd(n, s)$, and it follows from (5-3) that the stabilizer of α in this group is trivial; hence the orbit of α has cardinality $n/\gcd(n, s)$. This completes the proof of (2). □

Lemma 5.4. *Let $w = (f, \pi) \in \mathcal{W}$. Suppose that $i, j \in \Omega$ are such that $w(C_i) = C_j$, $w(C_j) = C_i$, and $w^2(\alpha) = \alpha$ for every $\alpha \in C_i \cup C_j$. Then there exists $0 \le s < n$ such that $w(\alpha) = \rho_i^{-s}\pi\rho_i^s(\alpha)$ for every $\alpha \in C_i \cup C_j$.*

Proof. Let $w(0, i) = (\bar{s}, j)$ and $w(0, j) = (\bar{t}, i)$ with $0 \le s, t < n$. From (5-3) and the fact that w commutes with ρ it follows that for every integer k we have $w(\bar{k}, i) = (\bar{s} + \bar{k}, j)$ and $w(\bar{k}, j) = (\bar{t} + \bar{k}, i)$. Using this we calculate $w^2(0, i) = w(\bar{s}, j) = (\bar{t} + \bar{s}, i)$. Since $w^2(0, i) = (0, i)$, this implies that $\bar{t} = -\bar{s}$; thus, for every integer k we have

$$w(\bar{k}, i) = (\bar{k} + \bar{s}, j) \quad \text{and} \quad w(\bar{k}, j) = (\bar{k} - \bar{s}, i). \tag{5-4}$$

Since $w(C_i) = C_j$ and $w(C_j) = C_i$, Lemma 5.3 implies that $\pi(i) = j$ and $\pi(j) = i$. It follows that for every $a \in A$ we have $\pi(a, i) = (a, j)$ and $\pi(a, j) = (a, i)$. Let $w' = \rho_i^{-s}\pi\rho_i^s$. If $\alpha = (\bar{k}, i) \in C_i$, then a simple calculation shows that $w'(\alpha) = (\bar{k} + \bar{s}, j)$, so $w'(\alpha) = w(\alpha)$ by (5-4). Similarly, if $\alpha = (\bar{k}, j) \in C_j$, then $w'(\alpha) = (\bar{k} - \bar{s}, i) = w(\alpha)$. Therefore $w(\alpha) = \rho_i^{-s}\pi\rho_i^s(\alpha)$ for every $\alpha \in C_i \cup C_j$. □

We can now prove the main result of this section.

Proposition 5.5. *Let $w \in \mathcal{W}$ and let $\mathcal{C}$ be the centralizer of w in $\mathcal{W}$.*

(1) *Suppose that w has cycle type $(2, D/2)$. Then the following hold*:
   (a) *Assume that $w(C_i) \ne C_i$ for all $i \in \Omega$. Then r is even, w is conjugate to the permutation $(1, 2)(3, 4)\cdots(r-1, r) \in S_r$, and $|\mathcal{C}| = (r/2)!(2n)^{r/2}$.*
   (b) *Assume $w(C_i) = C_i$ for some $i \in \Omega$. Then n is even and there exists $0 < \ell \le r$ such that $r - \ell$ is even and w is conjugate to the element $(\rho_1 \cdots \rho_\ell)^{n/2}\varepsilon$, where $\varepsilon = (\ell+1, \ell+2)\cdots(r-1, r) \in S_r$. Moreover, we have $|\mathcal{C}| = \ell!((r-\ell)/2)!n^\ell(2n)^{(r-\ell)/2}$.*

(2) *Suppose that w has cycle type $(2, n)$. Then the following hold*:
   (a) *Assume $w(C_i) = C_i$ for all $i \in \Omega$. Then n is even, there exist indices $i < j \in \Omega$ such that $w = (\rho_i\rho_j)^{n/2}$, and $|\mathcal{C}| = 2(r-2)!n^r$.*
   (b) *Assume $w(C_i) \ne C_i$ for some $i \in \Omega$. Then there exist indices $i < j \in \Omega$ and an integer $0 \le s < n$ such that $w = \rho_i^{-s}\tau\rho_i^s$, where $\tau = (i, j) \in S_r$. In this case, $|\mathcal{C}| = 2(r-2)!n^{r-1}$.*

(3) *Suppose that w moves exactly n elements of X. Then $w = \rho_i^s$ for some $i \in \Omega$ and some integer $0 < s < n$. Moreover, $|\mathcal{C}| = (r-1)!n^r$.*

Proof. Let $f \in A^r$ and $\pi \in S_r$ be such that $w = (f, \pi)$, and let (t_{ak}) be the type of w. We begin by proving 1(a). The hypothesis in (1) together with the fact that $\mathcal{W}$ acts faithfully on X imply that $w^2 = (f + f_\pi, \pi^2)$ is the identity element $(e, 1)$; in particular, $\pi^2 = 1$. Moreover, by Lemma 5.3 we have $w(C_i) = C_{\pi(i)}$ for every $i \in \Omega$, so $\pi(i) \ne i$ for every i. Hence the π-orbit of every element of Ω has cardinality 2. It follows that r is even, say $r = 2m$, and π is a product of m disjoint transpositions. We can now determine the type of w.

Let $\{i_1, \pi(i_1)\}, \ldots, \{i_m, \pi(i_m)\}$ be the orbits of π. Since π has no k-cycles if $k = 1$ or $k > 2$, then $t_{ak} = 0$ for all such k. When $k = 2$, t_{ak} is the number of indices $1 \le v \le m$ such that $f(i_v) + f(\pi(i_v)) = a$. Since

$\pi^2 = 1$, this is equivalent to $f(i_v) + f_\pi(i_v) = a$. Now, as mentioned above, $w^2 = (f + f_\pi, \pi^2) = (e, 1)$, so $f + f_\pi = e$ and therefore $f(i) + f_\pi(i) = 0$ for every $i \in \Omega$. Hence, the condition $f(i_v) + f_\pi(i_v) = a$ is equivalent to $a = 0$. Thus we have $t_{a2} = 0$ if $a \neq 0$, and $t_{02} = m$. This determines the type of w. It is now trivial to check that w has the same type as the permutation $\tau = (1, 2)(3, 4) \cdots (r - 1, r) \in S_r$. It follows from Proposition 5.2 that w is conjugate to τ and that $|\mathcal{C}| = m!(2n)^m$; this completes the proof of 1(a).

Next we prove 1(b). Suppose that $i \in \Omega$ satisfies $w(C_i) = C_i$. By Lemma 5.3, there exists $0 \leq s < n$ such that w acts like ρ_i^s on C_i, and the w-orbit of every element of C_i has cardinality $n/\gcd(n, s)$. By hypothesis every orbit has size 2, so $n/\gcd(n, s) = 2$, and hence n must be even and $s = n/2$.

Let $i_1, \ldots, i_\ell$ be all the indices i in Ω such that $w(C_i) = C_i$. Clearly, $0 < \ell \leq r$. Arguing as in the proof of 1(a), we see that π fixes i_k for each k, and that if $i \in \Omega \setminus \{i_1, \ldots, i_\ell\}$, then the orbit of i under π has size 2. This implies that $r - \ell$ is even, say $r - \ell = 2q$, and the disjoint cycle decomposition of π is a product of ℓ 1-cycles and q transpositions. The type of w is now easy to determine as done in case 1(a).

Clearly, $t_{ak} = 0$ if $k > 2$. Let $\{i_1\}, \ldots, \{i_\ell\}, \{j_1, \pi(j_1)\}, \ldots, \{j_q, \pi(j_q)\}$ be the orbits of π. Then t_{a2} is the number of indices $1 \leq v \leq q$ such that $f(j_v) + f_\pi(j_v) = a$. But $f + f_\pi = e$, so $t_{a2} = 0$ if $a \neq 0$, and $t_{02} = q$. To determine t_{a1} we need an additional observation. We know that for every index $1 \leq v \leq \ell$, w acts like $\rho_{i_v}^s$ on C_{i_v}. In particular, by (5-3) we have $w(0, i_v) = (\bar{s}, i_v)$. However, by (5-2), $w(0, i_v) = (f(i_v), i_v)$. Thus $f(i_v) = \bar{s}$ for all v. Now, t_{a1} is the number of indices $1 \leq v \leq \ell$ such that $f(i_v) = a$. Clearly then, $t_{a1} = 0$ if $a \neq \bar{s}$ and $t_{\bar{s}1} = \ell$. This determines the type of w.

Proposition 5.2 yields

$$|\mathcal{C}| = \ell! q! n^\ell (2n)^q.$$

Let $w' = (\rho_1 \cdots \rho_\ell)^s \varepsilon$, where $\varepsilon = (\ell + 1, \ell + 2) \cdots (r - 1, r) \in S_r$. A straightforward calculation shows that w' has the same type as w, and is therefore conjugate to w. This completes the proof of 1(b).

We now prove 2(a). If w acts nontrivially on m of the sets C_i, then the number of elements moved by w is mn; hence $m = 2$, so w acts trivially on all but two of these sets, say C_i and C_j with $i < j$. By Lemma 5.3, there exist integers $0 < u, v < n$ such that w acts like ρ_i^u on C_i and like ρ_j^v on C_j. The w-orbit of every element of C_i then has size $n/\gcd(n, u) = 2$, so n is even and $u = n/2$. Similarly, $v = n/2$. Thus w acts like $(\rho_i \rho_j)^{n/2}$ on all of X, and therefore $w = (\rho_i \rho_j)^{n/2}$. Letting $s = n/2$, we have $w = (s\delta_i + s\delta_j, 1)$; the type of w is now easily determined.

We have $t_{ak} = 0$ if $k > 1$, and t_{a1} is the number of indices $k \in \Omega$ such that $s\delta_i(k) + s\delta_j(k) = a$. Now, note that

$$s\delta_i(k) + s\delta_j(k) = 0, \quad \text{if } k \neq i, j, \quad \text{and} \quad s\delta_i(k) + s\delta_j(k) = \bar{s}, \quad \text{if } k = i \text{ or } j.$$

Hence $t_{a1} = 0$ if $a \notin \{0, \bar{s}\}$, $t_{\bar{s}1} = 2$, and $t_{01} = r - 2$. Proposition 5.2 now yields $|\mathcal{C}| = 2(r - 2)! n^r$; this proves 2(a).

Next we prove 2(b). By Lemma 5.3 we have $w(C_i) = C_j$ for some $j \neq i$. Then $w(C_j)$ must equal C_i, for otherwise w would move more than $2n$ elements of X. Thus $w(C_i) = C_j$, $w(C_j) = C_i$, and w acts trivially on C_k for all $k \neq i, j$. It follows from Lemma 5.3 that $\pi = (i, j)$. Reversing the roles of i and j if

necessary, we may assume that $i < j$. By Lemma 5.4, there exists $0 \le s < n$ such that $w(\alpha) = \rho_i^{-s}\pi\rho_i^s(\alpha)$ for every $\alpha \in C_i \cup C_j$. Clearly, this equality also holds if $\alpha \in C_k$ with $k \notin \{i, j\}$, so $w = \rho_i^{-s}\pi\rho_i^s$. We can now determine the type of w.

Since w and $\pi = (i, j)$ are conjugate, they have the same type. We thus find that $t_{ak} = 0$ if $k > 2$; $t_{a2} = 0$ if $a \neq 0$, and $t_{02} = 1$; $t_{a1} = 0$ if $a \neq 0$, and $t_{01} = r - 2$. Proposition 5.2 now yields $|\mathcal{C}| = 2(r-2)!n^{r-1}$; this completes the proof of 2(b).

Finally, we prove (3). It is easy to see that the n elements moved by w must form one of the sets C_i. This implies that $w(C_i) = C_i$ and w acts trivially on C_j for all $j \neq i$. By Lemma 5.3, there exists $0 < s < n$ such that $w(\alpha) = \rho_i^s(\alpha)$ for every $\alpha \in C_i$. This equality clearly holds for $\alpha \notin C_i$ as well, so $w = \rho_i^s$. Using the relation $w = \rho_i^s = (s\delta_i, 1)$, it is now a simple calculation to show that $t_{ak} = 0$ if $k > 1$, $t_{a1} = 0$ if $a \notin \{0, \bar{s}\}$, $t_{\bar{s}1} = 1$, and $t_{01} = r - 1$. Proposition 5.2 now yields $|\mathcal{C}| = (r-1)!n^r$. □

Having developed all of the necessary tools, we proceed to prove the main results of this article.

6. Genus computations for $n = 5$ and 6

Recall the following notation from Section 4: $K = \overline{\mathbb{Q}}(t)$, N/K is a splitting field of Φ_n, $G = \mathrm{Gal}(N/K)$, F is a subfield of N obtained by adjoining one root of Φ_n to K, and $\mathbb{P} = \{p_\infty\} \cup \{p_b \mid b \in \bigcup_{d\,|\,n} R_{n,d}\}$ is the set of places of K that ramify in F. Finally, for any intermediate field L in the extension N/K and any place p of K, $\mathbb{P}_L(p)$ denotes the set of places of L lying over p.

We begin this section by discussing an approach to the problem of computing the genera of subextensions of N/K. Let H be a subgroup of G with fixed field L, and let $g(L)$ denote the genus of L. We claim that if p is a place of K which ramifies in L, then $p \in \mathbb{P}$. Indeed, if p ramifies in L, then it ramifies in N. Letting $\mathfrak{P}$ be a place of N lying over p, the inertia group $I_{\mathfrak{P}\,|\,p}$ is nontrivial, so Corollary 3.5 implies that p ramifies in F. Hence $p \in \mathbb{P}$.

The Hurwitz genus formula [Stichtenoth 2009, Corollary 3.5.6] now yields

$$2g(L) - 2 = (-2)|G : H| + \sum_{p \in \mathbb{P}} \sum_{\mathfrak{q} \in \mathbb{P}_L(p)} (e_{\mathfrak{q}\,|\,p} - 1). \tag{6-1}$$

Let us define

$$g_{n,\infty}(H) = \sum_{\mathfrak{q} \in \mathbb{P}_L(p_\infty)} (e_{\mathfrak{q}\,|\,p_\infty} - 1),$$

and for every divisor d of n,

$$g_{n,d}(H) = \sum_{b \in R_{n,d}} \sum_{\mathfrak{q} \in \mathbb{P}_L(p_b)} (e_{\mathfrak{q}\,|\,p_b} - 1).$$

By (6-1) we have the following expression for the genus of L:

$$g(L) = 1 - |G : H| + \tfrac{1}{2}\left(g_{n,\infty}(H) + \sum_{d\,|\,n} g_{n,d}(H)\right). \tag{6-2}$$

The problem of computing $g(L)$ is thus reduced to the following: given any place $p \in \mathbb{P}$, compute the ramification index $e_{\mathfrak{q}\,|\,p}$ for every $\mathfrak{q} \in \mathbb{P}_L(p)$. Our method for doing this is based on the following lemma.

Lemma 6.1. *Let $p \in \mathbb{P}$, $\mathfrak{P} \in \mathbb{P}_N(p)$, and $I = I_{\mathfrak{P}\,|\,p}$. Let $\sigma_1, \ldots, \sigma_m$ be representatives of the distinct double cosets in $I\backslash G/H$. Then*

$$\{e_{\mathfrak{q}\,|\,p} : \mathfrak{q} \in \mathbb{P}_L(p)\} = \{|I^{\sigma_i} : I^{\sigma_i} \cap H| : 1 \le i \le m\}.$$

Proof. Since K is a function field over $\overline{\mathbb{Q}}$, we have $f_{\mathfrak{P}\,|\,p} = 1$ and therefore $D_{\mathfrak{P}\,|\,p} = I_{\mathfrak{P}\,|\,p} = I$. Using Lemma 3.1 we see that the set $\mathbb{P}_L(p)$ consists of the places $\sigma_i(\mathfrak{P}) \cap L$; moreover, if $\mathfrak{q} = \sigma_i(\mathfrak{P}) \cap L$, then $e_{\mathfrak{q}\,|\,p} = |I^{\sigma_i} : I^{\sigma_i} \cap H|$. The result follows immediately. □

For purposes of explicit computation it is convenient to use the isomorphisms $G \equiv \mathcal{W} \equiv \mathcal{Z}$ proved in Sections 5B–5D. With notation as in Lemma 6.1, suppose that one is able to identify the subgroup of $\mathcal{W}$ (or $\mathcal{Z}$) which corresponds to the inertia group I. It is then a finite computation to determine representatives $\sigma_1, \ldots, \sigma_m$ and to compute the indices $|I^{\sigma_i} : I^{\sigma_i} \cap H|$. Carrying out this calculation for every $p \in \mathbb{P}$, one obtains all the data needed to determine the numbers $g_{n,\infty}$ and $g_{n,d}$, and hence the genus of L.

The remainder of this section is devoted to showing that when $n = 5$ or 6 it is possible — and computationally feasible — to identify inertia groups $I_{\mathfrak{P}\,|\,p}$ for every $p \in \mathbb{P}$, and thus to compute the genus of any intermediate field in the extension N/K. In particular, this allows us to obtain the genera of the fixed fields of all the maximal subgroup of G, and by applying Proposition 2.4, to show that the sets E_5 and E_6 are finite.

In order to carry out all the necessary computations we have used version 2.23-1 of MAGMA [Bosma et al. 1997] running on a MacBook Pro with a 2.7 GHz Intel Core i5 processor and 8 GB of memory. The interested reader can find the code for our computations in [Krumm 2018a]. The code relies primarily on four intrinsic MAGMA functions: `WreathProduct`, `MaximalSubgroups`, `DoubleCosetRepresentatives`, and `meet`. The first function applied to $\mathbb{Z}/n\mathbb{Z}$ and S_r constructs the group $\mathcal{W}$ together with the natural embeddings $S_r \hookrightarrow \mathcal{W}$ and $(\mathbb{Z}/n\mathbb{Z})^r \hookrightarrow \mathcal{W}$. (It should be noted, however, that internally $\mathcal{W}$ is constructed as the group $\mathcal{Z}$.) Once $\mathcal{W}$ is constructed, the second function can be used to obtain the maximal subgroups of $\mathcal{W}$ up to conjugacy; the algorithm used is described in [Cannon and Holt 2004]. Given subgroups I and H of $\mathcal{W}$, the third function computes representatives of the double cosets in $I\backslash\mathcal{W}/H$. Finally, the fourth function can be used to compute the intersection of two subgroups of $\mathcal{W}$; the algorithm uses a backtrack method described in [Leon 1997].

Throughout this section we use the following notation. For $1 \le i \le r$ we let ρ_i be the element of $\mathcal{W}$ defined in Section 5F. As an automorphism of the graph Γ, ρ_i is a $1/n$ rotation of the i-th cycle. As an element of the group $\mathcal{Z}$, ρ_i is the i-th cycle in the decomposition of the permutation σ defined in Section 5B. For distinct indices $1 \le i, j \le r$ we let $\tau_{i,j}$ be the transposition $(i, j) \in S_r$ regarded as an element of $\mathcal{W}$. As an automorphism of Γ, $\tau_{i,j}$ interchanges the i-th and j-th cycles without performing any rotations.

Lemma 6.2. *The elements $\rho_1, \ldots, \rho_r$ are conjugate in $\mathcal{W}$. Moreover, if $i, j, u, v \in \{1, \ldots, r\}$ with $i \ne j$ and $u \ne v$, then $\rho_i\rho_j$ is conjugate to $\rho_u\rho_v$.*

Proof. This follows from Proposition 5.2. The type (t_{ak}) of ρ_i is independent of i; indeed, we have $t_{ak} = 0$ if $k > 1$, $t_{a1} = 0$ if $a \neq 0, 1$, $t_{01} = r - 1$, and $t_{11} = 1$. Similarly, if $i \neq j$, then the type (t_{ak}) of $\rho_i \rho_j$ independent of i and j: we have $t_{ak} = 0$ if $k > 1$, $t_{a1} = 0$ if $a \neq 0, 1$, $t_{01} = r - 2$, and $t_{11} = 2$. □

6A. ***The case n = 5.*** The polynomial Φ_5 has $D = 2\nu(5) = 30$ roots which can be partitioned into $r = D/5 = 6$ cycles. Hence, the graph Γ consists of six 5-cycles. The group $\mathcal{W}$ is $(\mathbb{Z}/5\mathbb{Z}) \wr S_6$, so $|G| = 5^6 6! = 11{,}250{,}000$. The set of places of K which ramify in F is

$$\mathbb{P} = \{p_\infty\} \cup \{p_b \mid b \in R_{5,5} \cup R_{5,1}\};$$

using Lemma 4.1 we obtain $\#R_{5,5} = 11$ and $\#R_{5,1} = 4$. We will henceforth identify G and $\mathcal{W}$ using the isomorphism $G \equiv \mathcal{W}$, where G acts on the roots of Φ_5 and $\mathcal{W}$ acts on the set $X = (\mathbb{Z}/5\mathbb{Z}) \times \{1, \ldots, 6\}$.

We define three subgroups of $\mathcal{W}$ by $A = \langle \tau_{1,2}\tau_{3,4}\tau_{5,6} \rangle$, $B = \langle \tau_{1,2} \rangle$, $C = \langle \rho_1 \rangle$.

Lemma 6.3. *Up to conjugation, A is the only subgroup of $\mathcal{W}$ generated by an element with cycle type* $(2, 15)$; *similarly, B is uniquely determined by the cycle type* $(2, 5)$, *and C by the cycle type* $(5, 1)$.

Proof. Suppose that $\tilde{A}$ is a subgroup of $\mathcal{W}$ generated by an element w with cycle type $(2, 15)$. We are then in the context of case 1 of Proposition 5.5. Moreover, since $n = 5$ is odd, case 1(b) is ruled out. Hence, by case 1(a), w is conjugate to $\tau_{1,2}\tau_{3,4}\tau_{5,6}$, and therefore $\tilde{A}$ is conjugate to A.

Now suppose that a subgroup $\tilde{B}$ is generated by an element w with cycle type $(2, 5)$. By case 2(b) of Proposition 5.5, w is conjugate to $\tau_{i,j}$ for some indices i, j. Clearly the permutations (i, j) and $(1, 2)$ are conjugates in S_6, so $\tau_{i,j}$ is conjugate to $\tau_{1,2}$ and therefore $\tilde{B}$ is conjugate to B.

Finally, suppose that a subgroup $\tilde{C}$ is generated by an element w with cycle type $(5, 1)$. By case 3 of Proposition 5.5, we have $w = \rho_i^s$ for some i and $0 < s < 5$. Note that $\langle \rho_i^s \rangle = \langle \rho_i \rangle$ since $|\rho_i| = 5$. By Lemma 6.2, w is conjugate to ρ_1^s, and therefore $\tilde{C} = \langle w \rangle$ is conjugate to $\langle \rho_1^s \rangle = \langle \rho_1 \rangle = C$. □

Lemma 6.4. (1) *There exists $\mathfrak{P} \in \mathbb{P}_N(p_\infty)$ such that $I_{\mathfrak{P}\,|\,p_\infty} = A$.*

(2) *For every $b \in R_{5,5}$ there exists $\mathfrak{P} \in \mathbb{P}_N(p_b)$ such that $I_{\mathfrak{P}\,|\,p_b} = B$.*

(3) *For every $b \in R_{5,1}$ there exists $\mathfrak{P} \in \mathbb{P}_N(p_b)$ such that $I_{\mathfrak{P}\,|\,p_b} = C$.*

Proof. Let $\mathfrak{P} \in \mathbb{P}_N(p_\infty)$. By Proposition 4.3 we have $I_{\mathfrak{P}\,|\,p_\infty} = \langle w \rangle$, where $w \in \mathcal{W}$ has cycle type $(2, 15)$. Thus, by Lemma 6.3, $I_{\mathfrak{P}\,|\,p_\infty}$ is conjugate to A. Replacing $\mathfrak{P}$ by a conjugate place if necessary, we then have $I_{\mathfrak{P}\,|\,p_\infty} = A$. This proves (1); the proofs of (2) and (3) are similar. □

Proposition 6.5. *Let H be a subgroup of $\mathcal{W}$ with fixed field L. Suppose that $\alpha_1, \ldots, \alpha_t$ are double coset representatives for $A\backslash\mathcal{W}/H$, $\beta_1, \ldots, \beta_u$ are representatives for $B\backslash\mathcal{W}/H$, and $\gamma_1, \ldots, \gamma_v$ are representatives for $C\backslash\mathcal{W}/H$. Then the genus of L is given by*

$$g(L) = 1 - |\mathcal{W} : H| + \tfrac{1}{2}(g_{5,\infty}(H) + g_{5,5}(H) + g_{5,1}(H)),$$

where

$$g_{5,\infty}(H) = \sum_{i=1}^{t} (|A^{\alpha_i} : A^{\alpha_i} \cap H| - 1), \tag{6-3}$$

$$g_{5,5}(H) = 11 \cdot \sum_{i=1}^{u} (|B^{\beta_i} : B^{\beta_i} \cap H| - 1), \tag{6-4}$$

$$g_{5,1}(H) = 4 \cdot \sum_{i=1}^{v} (|C^{\gamma_i} : C^{\gamma_i} \cap H| - 1). \tag{6-5}$$

Proof. The formula for $g(L)$ follows from (6-2). Let $p = p_\infty$. By Lemma 6.4, there exists $\mathfrak{P} \in \mathbb{P}_N(p)$ such that $I_{\mathfrak{P}\,|\,p} = A$. By Lemma 6.1 we have

$$\{e_{\mathfrak{q}\,|\,p} : \mathfrak{q} \in \mathbb{P}_L(p)\} = \{|A^{\alpha_i} : A^{\alpha_i} \cap H| : 1 \le i \le t\},$$

which implies (6-3). Now suppose that $b \in R_{5,5}$ and let $p = p_b$. By Lemma 6.4, there exists $\mathfrak{P} \in \mathbb{P}_N(p)$ such that $I_{\mathfrak{P}\,|\,p} = B$. Thus, by Lemma 6.1,

$$\{e_{\mathfrak{q}\,|\,p} : \mathfrak{q} \in \mathbb{P}_L(p)\} = \{|B^{\beta_i} : B^{\beta_i} \cap H| : 1 \le i \le u\},$$

and therefore

$$\sum_{\mathfrak{q} \in \mathbb{P}_L(p)} (e_{\mathfrak{q}\,|\,p} - 1) = \sum_{i=1}^{u} (|B^{\beta_i} : B^{\beta_i} \cap H| - 1).$$

Since the value of this sum is independent of b, and $\#R_{5,5} = 11$, then

$$g_{5,5}(H) = \sum_{b \in R_{5,5}} \sum_{\mathfrak{q} \in \mathbb{P}_L(p_b)} (e_{\mathfrak{q}\,|\,p_b} - 1) = 11 \cdot \sum_{i=1}^{u} (|B^{\beta_i} : B^{\beta_i} \cap H| - 1),$$

which proves (6-4). The proof of (6-5) is similar. □

We can now begin to prove Theorem 1.4.

Theorem 6.6. *The set E_5 is finite.*

Proof. Computing representatives for the conjugacy classes of maximal subgroups of $\mathcal{W}$, we obtain 8 subgroups which we denote by $M_1, \ldots, M_8$. The indices of these subgroups in $\mathcal{W}$ are given, respectively, by

$$|\mathcal{W} : M_i| : 3125, 15, 15, 10, 6, 6, 5, 2.$$

Let L_i be the fixed field of M_i. Fixing an index i, we may compute representatives for the double cosets in $A\backslash\mathcal{W}/M_i$, $B\backslash\mathcal{W}/M_i$, and $C\backslash\mathcal{W}/M_i$. The genus of L_i can then be obtained by applying Proposition 6.5. Carrying out these computations for $i = 1, \ldots, 8$ we obtain, respectively, the genera

$$9526, 21, 11, 9, 2, 12, 4, 5.$$

	M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8
$g_{5,\infty}$	1550	4	6	3	3	1	0	1
$g_{5,5}$	13750	66	44	33	11	33	0	11
$g_{5,1}$	10000	0	0	0	0	0	16	0

Table 1. Ramification data for the maximal subgroups of $\mathcal{W}$.

The result now follows from Proposition 2.4. The values of $g_{5,\infty}(M_i)$, $g_{5,5}(M_i)$, and $g_{5,1}(M_i)$ are shown in Table 1. □

6B. ***The case n = 6.*** Our next objective is to show that the set E_6 is finite. The structure of the proof is similar to the case $n = 5$, though the process of identifying the necessary inertia groups requires an additional step that was not present in that case.

The polynomial Φ_6 has $D = 2\nu(6) = 54$ roots which can be partitioned into $r = D/6 = 9$ cycles. Hence, the graph Γ consists of nine 6-cycles. The group $\mathcal{W}$ is $(\mathbb{Z}/6\mathbb{Z}) \wr S_9$, so $|G| = 6^9 9! = 3{,}656{,}994{,}324{,}480$. The set of places of K which ramify in F is $\mathbb{P} = \{p_\infty\} \cup \left\{p_b \mid b \in \bigcup_{d\,|\,6} R_{6,d}\right\}$. Using Lemma 4.1 we find that

$$\#R_{6,6} = 20, \quad \#R_{6,3} = 3, \quad \#R_{6,2} = 2, \quad \#R_{6,1} = 2. \tag{6-6}$$

We define several cyclic subgroups of $\mathcal{W}$. For $0 \le j \le 4$, let

$$\gamma_j = \left(\prod_{i=1}^{9-2j} \rho_i^3\right)\left(\prod_{i=0}^{j-1} \tau_{8-2i,9-2i}\right) \quad \text{and} \quad A_j = \langle \gamma_j \rangle.$$

In addition, let $B_0 = \langle \rho_1^3 \rho_2^3 \rangle$, $B_1 = \langle \tau_{1,2} \rangle$, $C = \langle \rho_1^3 \rangle$, $D = \langle \rho_1^2 \rangle$, $E = \langle \rho_1 \rangle$.

Lemma 6.7. *Up to conjugation, the groups A_j are the only subgroups of $\mathcal{W}$ generated by an element with cycle type* $(2, 27)$, *B_0 and B_1 are the only subgroups generated by an element with cycle type* $(2, 6)$, *and C, D, E are uniquely determined by the cycle types* $(2, 3)$, $(3, 2)$, *and* $(6, 1)$, *respectively.*

Proof. Suppose that $\tilde{A} = \langle w \rangle$, where $w \in \mathcal{W}$ has cycle type $(2, 27)$. We are then in the context of case 1 of Proposition 5.5. Moreover, since $r = 9$ is odd, case 1(b) must hold. Thus, there exists $0 < \ell \le 9$ such that $9 - \ell$ is even and w is conjugate to $v = (\rho_1 \cdots \rho_\ell)^3 (\tau_{\ell+1,\ell+2}) \cdots \tau_{8,9}$. Writing $9 - \ell = 2j$ with $0 \le j \le 4$, we have $v = \gamma_j$. Hence, $\tilde{A}$ is conjugate to A_j.

Suppose now that $\tilde{B} = \langle w \rangle$, where $w \in \mathcal{W}$ has cycle type $(2, 6)$. We are then in the context of case 2 of Proposition 5.5. In case 2(a) of the proposition, $w = (\rho_i \rho_j)^3$ for some indices $i \ne j$. By Lemma 6.2, this implies that w is conjugate to $(\rho_1 \rho_2)^3$, and therefore $\tilde{B}$ is conjugate to B_0. In case 2(b) of the proposition, w is conjugate to $\tau_{1,2}$ and $\tilde{B}$ is conjugate to B_1.

We now prove the uniqueness of the group C and omit the proofs for D and E, which are similar. Suppose that $\tilde{C} = \langle w \rangle$, where $w \in \mathcal{W}$ has cycle type $(2, 3)$. By case (3) of Proposition 5.5, we have $w = \rho_i^s$ with $1 \le i \le 9$ and $0 < s < 6$. In order for w to have cycle type $(2, 3)$ we must have $s = 3$; thus, by Lemma 6.2, w is conjugate to ρ_1^3 and $\tilde{C}$ is conjugate to C. □

Before continuing with the main discussion of this section, we prove a couple of auxiliary results. Returning to the general case of an arbitrary positive integer n, let θ be a root of Φ_n such that $F = K(\theta)$. Recall from Section 4 that F has an automorphism given by $\theta \mapsto \phi(\theta)$, and that F_0 denotes the fixed field of this automorphism.

Lemma 6.8. *Let* $\tau = \theta + \phi(\theta) + \cdots + \phi^{n-1}(\theta)$. *Then* $F_0 = K(\tau)$.

Proof. Following Morton [1996], we define the *trace* of a cycle in the graph Γ to be the sum of the elements in the cycle. Note that τ is the trace of the cycle containing θ. Let $P \in K[x]$ be the monic polynomial of degree r whose roots are the traces of all the cycles in Γ. By [Morton 1996, Corollary 3, p. 335], P is irreducible; hence P is the minimal polynomial of τ, and therefore $[K(\tau) : K] = r$. Clearly τ is fixed by ϕ, so $K(\tau) \subseteq F_0$. Now, since $[F : K] = D$ and $[F : F_0] = n$, then $[F_0 : K] = D/n = r = [K(\tau) : K]$. Thus $F_0 = K(\tau)$. □

We can now describe the subgroup of G corresponding to F_0.

Lemma 6.9. *Let* $\mathcal{O} = \{\theta, \phi(\theta), \ldots, \phi^{n-1}(\theta)\}$ *and let* H_0 *be the setwise stabilizer of* $\mathcal{O}$ *in* G. *Then* F_0 *is the fixed field of* H_0.

Proof. Let U and V be the subgroups of G defined by

$$U = \{\sigma \in G \mid \sigma(x) = x \text{ for every } x \in \mathcal{O}\} \quad \text{and} \quad V = \{\sigma \in G \mid \sigma(x) = x \text{ for every } x \in R \setminus \mathcal{O}\}.$$

A simple argument shows that $H_0 = UV$; see Example 2 in [Dummit and Foote 2004, p. 172].

The fact that ϕ is in the center of G implies that U is equal to the stabilizer of θ in G; thus $U = \mathrm{Gal}(N/F)$. It follows that F is the fixed field of U. Let L be the fixed field of H_0. Since $U \leq H_0$, then $L \subseteq F$. Defining τ as in Lemma 6.8, it is clear that τ is fixed by every element of H_0; hence $F_0 = K(\tau) \subseteq L$. We have thus shown that $F_0 \subseteq L \subseteq F$. To complete the proof we will show that $[F : L] = [F : F_0]$.

Identifying G with $\mathrm{Aut}(\Gamma)$ we see that V consists of the elements of G that act trivially on every cycle of Γ except possibly on the cycle containing θ. Thus the elements of V are the n rotations of the latter cycle, so $|V| = n$. By Galois theory we have $[F : L] = |UV|/|U| = |V|$, where the second equality uses the fact that $U \cap V = \{1\}$. We conclude that $[F : L] = n = [F : F_0]$. □

We return now to the case $n = 6$.

Lemma 6.10. (1) *There exists* $\mathfrak{P} \in \mathbb{P}_N(p_\infty)$ *such that* $I_{\mathfrak{P} \mid p_\infty} = A_4$.

(2) *For every* $b \in R_{6,6}$ *there exists* $\mathfrak{P} \in \mathbb{P}_N(p_b)$ *such that* $I_{\mathfrak{P} \mid p_b} = B_1$.

(3) *For every* $b \in R_{6,3}$ *there exists* $\mathfrak{P} \in \mathbb{P}_N(p_b)$ *such that* $I_{\mathfrak{P} \mid p_b} = C$.

(4) *For every* $b \in R_{6,2}$ *there exists* $\mathfrak{P} \in \mathbb{P}_N(p_b)$ *such that* $I_{\mathfrak{P} \mid p_b} = D$.

(5) *For every* $b \in R_{6,1}$ *there exists* $\mathfrak{P} \in \mathbb{P}_N(p_b)$ *such that* $I_{\mathfrak{P} \mid p_b} = E$.

Proof. Let $p = p_\infty$, $\mathfrak{P} \in \mathbb{P}_N(p)$, and $I = I_{\mathfrak{P} \mid p}$. By Proposition 4.3, I has a generator with cycle type $(2, 27)$. By Lemma 6.7, I must be conjugate to one of the groups A_j. Replacing $\mathfrak{P}$ by a conjugate ideal if necessary, we then have $I = A_j$ for some j. We claim that $I = A_4$.

To prove this we will use the number $S(p)$ defined in Proposition 4.4. By part (a) of the proposition, $S(p) = 9 - e_6 = 4$. We can calculate $S(p)$ in a different way by using the inertia group I as follows. Let H_0 be the subgroup of $\mathcal{W}$ defined in Lemma 6.9. Applying Lemma 6.1 we see that

$$S(p) = \sum_{i=1}^{m} (|I^{\sigma_i} : I^{\sigma_i} \cap H_0| - 1),$$

where $\sigma_1, \ldots, \sigma_m$ are double coset representatives for $I\backslash\mathcal{W}/H_0$. Assuming that $I = A_0, A_1, A_2, A_3, A_4$, respectively, we compute representatives σ_i and use the above formula to obtain $S(p) = 0, 1, 2, 3, 4$. However, we know that $S(p) = 4$, so necessarily $I = A_4$, as claimed. This proves (1). For the purposes of this computation, we identify $\mathcal{W}$ with the group $\mathcal{Z} \leq S_{54}$, so that H_0 is identified with the setwise stabilizer of the set $\{1, \ldots, 6\}$ in $\mathcal{Z}$. The code used for these computations is available in [Krumm 2018a].

Let $b \in R_{6,6}$, $p = p_b$, $\mathfrak{P} \in \mathbb{P}_N(p)$, and $I = I_{\mathfrak{P}\,|\,p}$. By Proposition 4.3, I has a generator with cycle type $(2, 6)$. By Lemma 6.7, I must be conjugate to either B_0 or B_1. Replacing $\mathfrak{P}$ by a conjugate ideal if necessary, we then have $I = B_0$ or B_1. We know that $S(p) = 1$ by part (b) of Proposition 4.4. Now, assuming that $I = B_0, B_1$, respectively, the above displayed formula yields $S(p) = 0, 1$; hence $I = B_1$. This proves (2).

Statements (3)-(5) follow easily from Proposition 4.3 and Lemma 6.7. □

Proposition 6.11. *Let H be a subgroup of $\mathcal{W}$ with fixed field L. For every group $I \in \{A_4, B_1, C, D, E\}$ let*

$$q_H(I) = \sum_{i=1}^{m} (|I^{\sigma_i} : I^{\sigma_i} \cap H| - 1),$$

where $\sigma_1, \ldots, \sigma_m$ are representatives of all the double cosets in $I\backslash\mathcal{W}/H$. Then the genus of L is given by

$$g(L) = 1 - |\mathcal{W} : H| + \tfrac{1}{2}(q_H(A_4) + 20q_H(B_1) + 3q_H(C) + 2q_H(D) + 2q_H(E)).$$

Proof. Let $p = p_\infty$. By Lemma 6.10, there exists $\mathfrak{P} \in \mathbb{P}_N(p)$ such that $I_{\mathfrak{P}\,|\,p} = A_4$. Using Lemma 6.1 we see that $g_{6,\infty}(H) = q_H(A_4)$. Now let $b \in R_{6,6}$, $p = p_b$, and let $\mathfrak{P} \in \mathbb{P}_N(p)$ satisfy $I_{\mathfrak{P}\,|\,p} = B_1$. By Lemma 6.1,

$$q_H(B_1) = \sum_{\mathfrak{q} \in \mathbb{P}_L(p)} (e_{\mathfrak{q}\,|\,p} - 1).$$

Since this holds for every $b \in R_{6,6}$, then (6-6) yields $g_{6,6}(H) = 20q_H(B_1)$. By a similar argument we show that

$$g_{6,3}(H) = 3q_H(C), \quad g_{6,2}(H) = 2q_H(D), \quad \text{and} \quad g_{6,1}(H) = 2q_H(E).$$

The stated formula for the genus of L is now a consequence of (6-1). □

We can now prove a second part of Theorem 1.4.

Theorem 6.12. *The set E_6 is finite.*

	M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	M_9	M_{10}	M_{11}
$q_{M_i}(A_4)$	416	132	120	60	40	16	4	0	1	0	1
$q_{M_i}(B_1)$	420	105	64	35	21	7	1	0	1	1	0
$q_{M_i}(C)$	0	0	128	0	0	0	0	0	1	0	1
$q_{M_i}(D)$	0	0	0	0	0	0	0	2	0	0	0
$q_{M_i}(E)$	0	0	128	0	0	0	0	2	1	0	1

Table 2. Ramification data for the maximal subgroups of $\mathcal{W}$.

Proof. Computing representatives for the conjugacy classes of maximal subgroups of $\mathcal{W}$, we obtain 11 subgroups which we denote by $M_1, \ldots, M_{11}$. The indices of these subgroups in $\mathcal{W}$ are given, respectively, by

$$|\mathcal{W} : M_i| : 840, 280, 256, 126, 84, 36, 9, 3, 2, 2, 2.$$

Let L_i be the fixed field of M_i. Fixing an index i, we may compute the numbers $q_{M_i}(I)$ for $I \in \{A_4, B_1, C, D, E\}$. The genus of L_i can then be obtained by applying Proposition 6.11. Carrying out these computations for $i = 1, \ldots, 11$ we obtain, respectively, the genera

$$3569, 837, 765, 255, 147, 43, 4, 2, 12, 9, 2.$$

By Proposition 2.4, this implies that E_6 is finite. The values of $q_{M_i}(I)$ are shown in Table 2. □

7. Genus bounds for $n > 6$

The methods used in the previous section for $n = 5$ and 6 can, in principle, be applied to higher values of n; however, there are computational limitations which make this impractical. Firstly, for $n > 10$ there are issues of both memory and time which prevent us from computing the maximal subgroups of $\mathcal{W}$. Thus, we are restricted to considering only $n = 7, 8, 9, 10$. Furthermore, even for these values of n there are similar complications in the crucial step of computing double coset representatives. Hence, it would appear that our methods cannot be extended beyond $n = 6$. However, a modification of the method will allow us to show that E_7 and E_9 are finite.

Recall that our main goal is to show that the genera of the function fields corresponding to maximal subgroups of G are all greater than 1. In the cases $n = 5, 6$ we did this by calculating the exact values of these genera, although it would be sufficient to prove a lower bound greater than 1. In this section we will show that, as long as the maximal subgroups of $\mathcal{W}$ can be computed, it is possible to obtain lower bounds for the required genera. In the cases $n = 7, 9$ these bounds will suffice to prove the desired result. Unfortunately, the bounds are not good enough when $n = 8, 10$; the difficulties are explained in Section 7B. We keep here all of the notation introduced in earlier sections.

Lemma 7.1. *Let H be a subgroup of G with fixed field L, and $a = |G : H|$. Let p be a place of K, let $\{\mathfrak{q}_1, \ldots, \mathfrak{q}_s\} = \mathbb{P}_L(p)$, and $e_i = e_{\mathfrak{q}_i \mid p}$. Suppose that u is an upper bound for the number of indices i such*

that $e_i = 1$. *Then*

$$\sum_{i=1}^{s}(e_i - 1) \geq \lceil a - \lfloor (u+a)/2 \rfloor \rceil.$$

Proof. Let x be the number of indices i such that $e_i = 1$, and let $y = s - x$. Note that $a = e_1 + \cdots + e_s \geq x + 2y$. Since $x \leq u$, this implies $x + y \leq (u+a)/2$. Thus $s \leq \lfloor (u+a)/2 \rfloor$ and therefore

$$\sum_{i=1}^{s}(e_i - 1) = a - s \geq a - \lfloor (u+a)/2 \rfloor,$$

from which the result follows immediately. □

7A. ***The case of odd n.*** Assume that n is odd. Using Lemma 7.1, we now explain how to obtain lower bounds for the genera of subextensions of N/K. Define subsets Θ_n and Λ_n of $\mathcal{W}$ by

$$\Theta_n = \{\rho_i^s \mid 1 \leq i \leq r, 0 < s < n\}, \quad \text{and} \quad \Lambda_n = \{\rho_i^{-s}\tau_{i,j}\rho_i^s \mid 1 \leq i < j \leq r, 0 \leq s < n\}.$$

For every subgroup H of $\mathcal{W}$ and every divisor d of n, let

$$u_{n,d}(H) = \begin{cases} (r-1)!n^r \#(H \cap \Theta_{n,d})/|H| & \text{if } d < n, \\ 2(r-2)!n^{r-1}\#(H \cap \Lambda_n)/|H| & \text{if } d = n, \end{cases}$$

and

$$g'_{n,d}(H) = (\deg \Delta_{n,d})\left\lceil |\mathcal{W} : H| - \left\lfloor \frac{u_{n,d}(H) + |\mathcal{W} : H|}{2} \right\rfloor \right\rceil.$$

Here, $\Theta_{n,d}$ denotes the set of elements of Θ_n having cycle type $(n/d, d)$.

Proposition 7.2. *With notation as above, let* L *be the fixed field of* H. *Then the genus of* L *satisfies*

$$g(L) \geq \left\lceil 1 - |\mathcal{W} : H| + \tfrac{1}{2}\sum_{d \mid n} \max(g'_{n,d}(H), 0) \right\rceil. \tag{7-1}$$

Proof. Let d be a proper divisor of n, and let $b \in R_{n,d}$. If $\mathfrak{P}$ is a place of N lying over p_b, Proposition 4.3 implies that the inertia group $I_{\mathfrak{P} \mid p_b}$ is generated by an element γ with cycle type $(n/d, d)$. By part (3) of Proposition 5.5, we have $\gamma = \rho_i^s$ with $1 \leq i \leq r$ and $0 < s < n$. Moreover, the order of the centralizer of γ is given by $|C_{\mathcal{W}}(\gamma)| = (r-1)!n^r$. Thus, by Corollary 3.5, the number of places $\mathfrak{q} \in \mathbb{P}_L(p_b)$ such that $e_{\mathfrak{q} \mid p_b} = 1$ is equal to

$$(r-1)!n^r s(H, \gamma)/|H|,$$

where $s(H, \gamma)$ is the number of conjugates of γ which belong to H. Note that every conjugate of γ belongs to $\Theta_{n,d}$, so that $s(H, \gamma) \leq \#(H \cap \Theta_{n,d})$. It follows that the number of places $\mathfrak{q} \in \mathbb{P}_L(p_b)$ such that $e_{\mathfrak{q} \mid p_b} = 1$ is bounded above by $u_{n,d}(H)$. Letting $a = |\mathcal{W} : H|$, Lemma 7.1 implies that

$$\sum_{\mathfrak{q} \in \mathbb{P}_L(p_b)} (e_{\mathfrak{q} \mid p_b} - 1) \geq \lceil a - \lfloor (u_{n,d}(H) + a)/2 \rfloor \rceil.$$

Recalling the number $g_{n,d}(H)$ defined in Section 6, the above inequality implies that $g_{n,d}(H) \geq g'_{n,d}(H)$ and therefore $g_{n,d}(H) \geq \max(g'_{n,d}(H), 0)$.

By a similar argument we can show that $g_{n,n}(H) \geq \max(g'_{n,n}(H), 0)$. Let $b \in R_{n,n}$ and $\mathfrak{P} \in \mathbb{P}_N(p_b)$. Then $I_{\mathfrak{P}\,|\,p_b} = \langle \gamma \rangle$, where γ has cycle type $(2, n)$. Since n is odd, part 2(b) of Proposition 5.5 implies that $\gamma = \rho_i^{-s} \tau_{i,j} \rho_i^s$ for some $1 \leq i < j \leq r$ and $0 \leq s < n$. Moreover, $|C_{\mathcal{W}}(\gamma)| = 2(r-2)!n^{r-1}$. The number of places $\mathfrak{q} \in \mathbb{P}_L(p_b)$ such that $e_{\mathfrak{q}\,|\,p_b} = 1$ is therefore given by

$$2(r-2)!n^{r-1} s(H, \gamma)/|H|.$$

Now, every conjugate of γ belongs to Λ_n, so $s(H, \gamma) \leq \#(H \cap \Lambda_n)$. The number of places $\mathfrak{q} \in \mathbb{P}_L(p_b)$ with $e_{\mathfrak{q}\,|\,p_b} = 1$ is thus bounded above by $u_{n,n}$. Letting $a = |\mathcal{W} : H|$, we have

$$\sum_{\mathfrak{q} \in \mathbb{P}_L(p_b)} (e_{\mathfrak{q}\,|\,p_b} - 1) \geq \left\lceil a - \lfloor (u_{n,n}(H) + a)/2 \rfloor \right\rceil,$$

which implies that $g_{n,n}(H) \geq g'_{n,n}(H)$. We have thus proved:

$$g_{n,d}(H) \geq \max(g'_{n,d}(H), 0), \quad \text{for every divisor } d \text{ of } n.$$

Now (7-1) follows from the genus formula (6-2). □

Remark 7.3. Note that in proving the bound (7-1) we have disregarded the contribution to the genus coming from ramified places lying over p_∞. Though the bound would certainly be improved if these places were considered, doing so would substantially increase the amount of time and memory required to compute the bound. In particular, it would require determining the intersection $H \cap C$, where C is the set of all conjugates in $\mathcal{W}$ of the permutation $(1, 2)(3, 4) \cdots (r-1, r)$. Now, part 1(a) of Proposition 5.5 implies that $\#C = (n^r r!)/((r/2)!(2n)^{r/2}) \geq n^{r/2}$, which suggests that C might be difficult to construct in practice. And indeed, our attempts to compute all the elements of C in the case $n = 7$ failed due to excessive memory requirements.

Remark 7.4. In order to compute the number on the right-hand side of (7-1), the key step is to determine the cardinalities of the sets $H \cap \Theta_{n,d}$ and $H \cap \Lambda_n$, which would be difficult to do if all the sets involved were quite large. Fortunately, while the group H may be extremely large (for instance, H might be the largest maximal subgroup of the Galois group of Φ_9, in which case $|H| \approx 9.73 \times 10^{127}$), the sets Λ_n and $\Theta_{n,d}$ are small. Indeed, $\#\Theta_{n,d} \leq \#\Theta_n = r(n-1)$ and $\#\Lambda_n = n \cdot \binom{r}{2}$. This makes it computationally feasible to construct the sets $H \cap \Lambda_n$ and $H \cap \Theta_{n,d}$, and hence to compute the desired lower bound.

We can now complete the proof of Theorem 1.4. The finiteness of E_7 and E_9 is proved by a series of computations carried out using MAGMA; the code used for these computations is available in [Krumm 2018a].

Theorem 7.5. *The sets E_7 and E_9 are finite.*

Proof. We consider first the case of E_7. The polynomial Φ_7 has $D = 126$ roots which can be partitioned into $r = 18$ cycles. Thus, $\mathcal{W} = (\mathbb{Z}/7\mathbb{Z}) \wr S_{18}$. Constructing the group $\mathcal{W}$ and computing representatives for

the conjugacy classes of maximal subgroups of $\mathcal{W}$, we obtain 16 groups which we denote by $M_1, \ldots, M_{16}$. The sets Θ_7 and Λ_7 are easily constructed; we find that $\#\Theta_7 = 108$ and $\#\Lambda_7 = 1071$.

Let L_i denote the fixed field of M_i. For each subgroup M_i we compute the numbers $u_{7,7}(M_i)$ and $u_{7,1}(M_i)$, and use these to calculate $g'_{7,7}(M_i)$ and $g'_{7,1}(M_i)$. This is a trivial computation given the small size of the sets Θ_7 and Λ_7. The inequality (7-1) then yields a lower bound for $g(L_i)$.

Carrying out these calculations, the lowest lower bound we obtain for the genera $g(L_i)$ is 6; hence $g(L_i) > 1$ for every i, which implies that E_7 is finite. The total time required for all of the above computations is 0.42 s.

The proof of finiteness of E_9 follows the same steps as above. In this case the lowest lower bound we obtain for $g(L_i)$ is 4. Total computation time is 197 s, with 179 s spent computing the maximal subgroups of $\mathcal{W}$. □

7B. ***The case of even n.*** In the case where n is even, a bound similar to (7-1) can be proved; indeed, this only requires modifying the definition of the number $u_{n,n}(H)$. Unfortunately, when $n = 8$ or 10 the bounds for the genera $g(L_i)$ obtained in this way are not greater than 1; in fact many of them are negative. We suspect, therefore, that most of the ramification in the extensions L_i/K occurs over the place p_∞. In order to improve the bounds for $g(L_i)$ we would have to determine the genus contribution coming from places lying over p_∞. However, as discussed in Remark 7.3, it is computationally infeasible to do this. Thus, we are unable to improve the bounds enough to show that E_8 and E_{10} are finite.

8. Density results

Having proved Theorem 1.4, we now turn our attention to Theorem 1.5. Recall that if n is a positive integer and $c \in \mathbb{Q}$, we denote by $T_{n,c}$ the set of prime numbers p such that the map $\phi_c(x) = x^2 + c$ does not have a point of period n in $\mathbb{Q}_p$. By applying Lemma 8.1 below we will be able to calculate the density of $T_{n,c}$ for $n \in \{5, 6, 7, 9\}$ and all but finitely many $c \in \mathbb{Q}$.

For every polynomial $F \in \mathbb{Q}[x]$, let S_F be the set of all primes p such that F has a root in $\mathbb{Q}_p$. The Chebotarev density theorem implies that the density of S_F, which we denote by $\delta(S_F)$, exists and can be computed if the Galois group of F is known. More precisely, we have the following result.

Lemma 8.1. *Let $F \in \mathbb{Q}[x]$ be a separable polynomial of degree $D \geq 1$. Let S be a splitting field for F, and set $G = \mathrm{Gal}(S/\mathbb{Q})$. Let $\alpha_1, \ldots, \alpha_D$ be the roots of F in S and, for each index i, let G_i denote the stabilizer of α_i under the action of G. Then the Dirichlet density of S_F is given by*

$$\delta(S_F) = \frac{\left|\bigcup_{i=1}^{D} G_i\right|}{|G|}. \tag{8-1}$$

Proof. This follows from Theorem 2.1 in [Krumm 2016]. □

Note that for the purpose of computing $\delta(S_F)$ using the formula (8-1), the group G may be replaced with any permutation group $\mathcal{G}$ such that $G \equiv \mathcal{G}$. Fixing a positive integer n, let G_n be the Galois group of Φ_n over $\mathbb{Q}(t)$ and let $\mathcal{G} = \mathrm{Aut}(\Gamma)$, where Γ is the graph defined in Section 5A. Recall that $G_n \equiv \mathcal{G}$.

Lemma 8.2. *Let $\mathcal{M}$ be the set of all elements of $\mathcal{G}$ having no fixed point. The cardinality of $\mathcal{M}$ is given by the formula*

$$\#\mathcal{M} = \sum_{i=0}^{r} (n-1)^i \cdot n^{r-i} \cdot d(r,i),$$

where

$$d(r,i) = \binom{r}{i}(r-i)! \sum_{k=0}^{r-i} \frac{(-1)^k}{k!}.$$

Proof. The number $d(r,i)$ counts the permutations in S_r which fix exactly i elements of the set $\{1,\dots,r\}$. The above formula for $d(r,i)$ is proved by an inclusion-exclusion argument; see Example 2.2.1 in [Stanley 2012].

For $0 \le i \le r$, let $\mathcal{M}_i$ be the set of elements of $\mathcal{M}$ which fix exactly i cycles of Γ. Clearly $\mathcal{M}$ is a disjoint union of the sets $\mathcal{M}_i$, so in order to prove the lemma it suffices to show that

$$\#\mathcal{M}_i = (n-1)^i \cdot n^{r-i} \cdot d(r,i).$$

Recall that every element $\sigma \in \mathcal{G}$ has a unique representation of the form $\rho_1^{a_1} \cdots \rho_r^{a_r} \pi$, where $\pi \in S_r$ describes the action of σ on the set of cycles of Γ, ρ_k represents a $(1/n)$ rotation on the k-th cycle, and $0 \le a_k < n$.

Let $0 \le i \le r$. Then an element $\sigma \in \mathcal{G}$ represented as above belongs to $\mathcal{M}_i$ if and only if there exist indices $k_1,\dots,k_i \in \{1,\dots,r\}$ such that π fixes $k_1,\dots,k_i$ and has no other fixed points; and $a_{k_j} > 0$ for $j = 1,\dots,i$. In constructing elements of $\mathcal{M}_i$ we therefore have $d(r,i)$ choices for π, $n-1$ choices for the exponents a_{k_j}, and n choices for the remaining $r-i$ exponents. It follows that $\#\mathcal{M}_i = (n-1)^i \cdot n^{r-i} \cdot d(r,i)$, as required. □

Proof of Theorem 1.5. Let $\Delta(t)$ be the discriminant of Φ_n and let

$$E = \{c \in \mathbb{Q} \mid \Delta(c) = 0\} \cup E_5 \cup E_6 \cup E_7 \cup E_9.$$

By the results of Sections 6 and 7, E is a finite set. Fix $n \in \{5,6,7,9\}$ and $c \in \mathbb{Q} \setminus E$. Since $c \notin E_n$, we have $G_{n,c} \cong G_n$. This implies that $G_{n,c} \equiv G_n$, where $G_{n,c}$ acts on the roots of $\Phi_n(c,x)$. Indeed, since $\Delta(c) \neq 0$, there is a subgroup H of G_n such that $G_{n,c} \equiv H$ (see Theorem 2.9 in [Lang 2002, Chapter VII]). By order considerations, H must be equal to G_n.

Let $S_{n,c}$ be the set of primes p such that $\Phi_n(c,x)$ has a root in $\mathbb{Q}_p$. The fact that $\Delta(c) \neq 0$ implies that $S_{n,c}$ is the complement of $T_{n,c}$. Indeed, every root of $\Phi_n(c,x)$ has period n under ϕ_c; see Theorem 2.4(c) in [Morton and Patel 1994].

Since $G_{n,c} \equiv G_n \equiv \mathcal{G}$, Lemma 8.1 applied to $F(x) = \Phi_n(c,x)$ yields

$$\delta(S_{n,c}) = \frac{\left|\bigcup_{\alpha\in\Gamma} \mathcal{G}_\alpha\right|}{|\mathcal{G}|},$$

where $\mathcal{G}_\alpha$ is the stabilizer of α in $\mathcal{G}$. It follows that $\delta(T_{n,c}) = (\#\mathcal{M})/|\mathcal{G}|$, where $\mathcal{M}$ is defined as in Lemma 8.2. Using this lemma we obtain

$$\begin{aligned}
\delta(T_{5,c}) &= \tfrac{9210721}{6!5^6} \approx 0.8187,\\
\delta(T_{6,c}) &= \tfrac{3095578863701}{9!6^9} \approx 0.8465,\\
\delta(T_{7,c}) &\approx 0.8669,\\
\delta(T_{9,c}) &\approx 0.8948.
\end{aligned}$$

This completes the proof of the theorem. □

9. The exceptional sets E_n

We end this article with a brief discussion concerning the elements of the sets E_n. Recall the following notation introduced in Section 2: S is a splitting field of Φ_n over $\mathbb{Q}(t)$, $G_n = \mathrm{Gal}(S/\mathbb{Q}(t))$, $M_1, \ldots, M_s$ are representatives of the conjugacy classes of maximal subgroups of G_n, and $\mathcal{X}$ is the smooth projective curve with function field S.

Our approach to proving the finiteness of E_n for $n > 4$ is based on Lemma 2.1, which shows that E_n is finite if every quotient curve $\mathcal{X}/M_i$ has genus greater than 1. The proof of the lemma suggests that we may determine the elements of E_n by finding a certain finite set $\mathcal{E}$ and determining all the rational points on the curves $\mathcal{X}/M_i$. The set $\mathcal{E}$ as well as affine models for these curves can be obtained using the methods of the article [Krumm and Sutherland 2017]; however, the rational points on $\mathcal{X}/M_i$ seem impossible to determine due to the large genera of the curves. (For instance, when $n = 5$ one of the curves has genus 9526, as seen in the proof of Theorem 6.6.) Hence, the problem of explicitly determining E_n seems intractable at present. Nevertheless, it is possible to prove some basic results about the elements of E_n.

Proposition 9.1. *For every positive integer n we have $\{0, -2\} \subseteq E_n$.*

Proof. For every $c \in \mathbb{Q}$, the polynomial $\Phi_n(c, x)$ divides $\phi_c^n(x) - x$, where $\phi_c(x) = x^2 + c$. In particular, $\Phi_n(0, x)$ divides $x^{2^n} - x$, which implies that $\Phi_n(0, x)$ splits over a cyclotomic field. It follows that the Galois group $G_{n,0}$ is abelian, hence not isomorphic to G_n, since $G_n \cong (\mathbb{Z}/n\mathbb{Z}) \wr S_r$. Thus $0 \in E_n$.

For $c = -2$ the polynomial ϕ_c is a Chebyshev polynomial satisfying

$$\phi_c(x + 1/x) = x^2 + 1/x^2.$$

We claim that the polynomial $\Phi_n(-2, x)$ splits over the cyclotomic field $\mathbb{Q}(\zeta)$, where ζ is a primitive $(2^{2n} - 1)$-th root of unity; as above, this will imply that $-2 \in E_n$. Suppose that $\alpha \in \overline{\mathbb{Q}}$ is a root of $\Phi_n(-2, x)$, and let $\beta \in \overline{\mathbb{Q}}$ satisfy $\beta + 1/\beta = \alpha$. Then $\beta^{2^n} + 1/\beta^{2^n} = \beta + 1/\beta$, which implies that $(\beta^{2^n+1} - 1)(\beta^{2^n-1} - 1) = 0$ and hence $\beta^{2^{2n}-1} = 1$. Thus β, and therefore α, belongs to $\mathbb{Q}(\zeta)$. This proves the claim. □

Given a positive integer n for which E_n is finite, one can attempt to find all the elements of E_n by carrying out an exhaustive search within specified height bounds. Recall that the height of a rational

number $\frac{a}{b}$ with $\gcd(a,b)=1$ is given by $\max(|a|,|b|)$. Fixing a height bound h, it a straightforward procedure to construct the set $B(h)$ of all rational numbers having height at most h. One can then construct all the polynomials $\Phi_n(c,x)$ for $c \in B(h)$, compute their Galois groups $G_{n,c}$ (for instance, using the algorithm of Fieker and Klüners [2014], which is implemented in MAGMA), and check whether $G_{n,c} \cong (\mathbb{Z}/n\mathbb{Z}) \wr S_r$. The cost of carrying out this computation grows quickly with n, given the large degree of Φ_n. For $n=7$ the degree of Φ_n is 126, and the above computation is very slow even for small height bounds h. However, for $n=5$ and 6 we have the following result.

Proposition 9.2. *Let $B(h)$ denote the set of all rational numbers with height at most h. Then*

$$E_5 \cap B(50) = \left\{-2, -\tfrac{16}{9}, -\tfrac{3}{2}, -\tfrac{4}{3}, -\tfrac{5}{8}, 0\right\} \quad \textit{and} \quad E_6 \cap B(20) = \{-4, -2, 0\}.$$

References

[Artin 2006] E. Artin, *Algebraic numbers and algebraic functions*, AMS Chelsea Publishing, Providence, RI, 2006. MR Zbl

[Beckmann 1994] S. Beckmann, "On finding elements in inertia groups by reduction modulo p", *J. Algebra* **164**:2 (1994), 415–429. MR Zbl

[Bosma et al. 1997] W. Bosma, J. Cannon, and C. Playoust, "The Magma algebra system, I: The user language", *J. Symbolic Comput.* **24**:3-4 (1997), 235–265. MR Zbl

[Bousch 1992] T. Bousch, *Sur quelques problèmes de dynamique holomorphe*, Ph.D. thesis, Université de Paris-Sud, Centre d'Orsay, 1992.

[Cannon and Holt 2004] J. Cannon and D. F. Holt, "Computing maximal subgroups of finite groups", *J. Symbolic Comput.* **37**:5 (2004), 589–609. MR Zbl

[Dixon and Mortimer 1996] J. D. Dixon and B. Mortimer, *Permutation groups*, Graduate Texts in Mathematics **163**, Springer, 1996. MR Zbl

[Dummit and Foote 2004] D. S. Dummit and R. M. Foote, *Abstract algebra*, 3rd ed., John Wiley & Sons, Hoboken, NJ, 2004. MR Zbl

[Efrat 2006] I. Efrat, *Valuations, orderings, and Milnor K-theory*, Mathematical Surveys and Monographs **124**, Amer. Math. Soc., Providence, RI, 2006. MR Zbl

[Faltings 1983] G. Faltings, "Endlichkeitssätze für abelsche Varietäten über Zahlkörpern", *Invent. Math.* **73**:3 (1983), 349–366. MR Zbl

[Fieker and Klüners 2014] C. Fieker and J. Klüners, "Computation of Galois groups of rational polynomials", *LMS J. Comput. Math.* **17**:1 (2014), 141–158. MR Zbl

[Flynn et al. 1997] E. V. Flynn, B. Poonen, and E. F. Schaefer, "Cycles of quadratic polynomials and rational points on a genus-2 curve", *Duke Math. J.* **90**:3 (1997), 435–463. MR Zbl

[Frucht 1949] R. Frucht, "On the groups of repeated graphs", *Bull. Amer. Math. Soc.* **55** (1949), 418–420. MR Zbl

[Harary 1969] F. Harary, *Graph theory*, Addison-Wesley Publishing Co., Reading, MA, 1969. MR Zbl

[Kerber 1971] A. Kerber, *Representations of permutation groups, I*, Lecture Notes in Mathematics **240**, Springer, 1971. MR Zbl

[Krumm 2016] D. Krumm, "A local-global principle in the dynamics of quadratic polynomials", *Int. J. Number Theory* **12**:8 (2016), 2265–2297. MR Zbl

[Krumm 2018a] D. Krumm, "code for the computations in the article "A finiteness theorem for specializations of dynatomic polynomials"", 2018, Available at https://github.com/davidkrumm/finiteness_dynatomic. Magma code.

[Krumm 2018b] D. Krumm, "Galois groups in a family of dynatomic polynomials", *J. Number Theory* **187** (2018), 469–511. MR Zbl

[Krumm and Sutherland 2017] D. Krumm and N. Sutherland, "Galois groups over rational function fields and explicit Hilbert irreducibility", preprint, 2017. arXiv

[Lang 2002] S. Lang, *Algebra*, 3rd ed., Graduate Texts in Mathematics **211**, Springer, 2002. MR Zbl

[Leon 1997] J. S. Leon, "Partitions, refinements, and permutation group computation", pp. 123–158 in *Groups and computation, II* (New Brunswick, NJ, 1995), edited by L. Finkelstein and W. M. Kantor, DIMACS Ser. Discrete Math. Theoret. Comput. Sci. **28**, Amer. Math. Soc., Providence, RI, 1997. MR Zbl

[Morton 1992] P. Morton, "Arithmetic properties of periodic points of quadratic maps", *Acta Arith.* **62**:4 (1992), 343–372. MR Zbl

[Morton 1996] P. Morton, "On certain algebraic curves related to polynomial maps", *Compositio Math.* **103**:3 (1996), 319–350. MR Zbl

[Morton 1998] P. Morton, "Arithmetic properties of periodic points of quadratic maps, II", *Acta Arith.* **87**:2 (1998), 89–102. MR Zbl

[Morton and Patel 1994] P. Morton and P. Patel, "The Galois theory of periodic points of polynomial maps", *Proc. London Math. Soc.* (3) **68**:2 (1994), 225–263. MR Zbl

[Morton and Vivaldi 1995] P. Morton and F. Vivaldi, "Bifurcations and discriminants for polynomial maps", *Nonlinearity* **8**:4 (1995), 571–584. MR Zbl

[Neukirch 1999] J. Neukirch, *Algebraic number theory*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences] **322**, Springer, 1999. MR Zbl

[Poonen 1998] B. Poonen, "The classification of rational preperiodic points of quadratic polynomials over $\mathbf{Q}$: a refined conjecture", *Math. Z.* **228**:1 (1998), 11–29. MR Zbl

[Rosen 2002] M. Rosen, *Number theory in function fields*, Graduate Texts in Mathematics **210**, Springer, 2002. MR Zbl

[Rotman 1995] J. J. Rotman, *An introduction to the theory of groups*, 4th ed., Graduate Texts in Mathematics **148**, Springer, 1995. MR Zbl

[Serre 2008] J.-P. Serre, *Topics in Galois theory*, 2nd ed., Research Notes in Mathematics **1**, A K Peters, Ltd., Wellesley, MA, 2008. MR

[Stanley 2012] R. P. Stanley, *Enumerative combinatorics, Volume 1*, 2nd ed., Cambridge Studies in Advanced Mathematics **49**, Cambridge University Press, 2012. MR Zbl

[Stichtenoth 2009] H. Stichtenoth, *Algebraic function fields and codes*, 2nd ed., Graduate Texts in Mathematics **254**, Springer, 2009. MR Zbl

[Stoll 2008] M. Stoll, "Rational 6-cycles under iteration of quadratic polynomials", *LMS J. Comput. Math.* **11** (2008), 367–380. MR Zbl

Communicated by Joseph H. Silverman
Received 2018-05-28 Revised 2019-01-22 Accepted 2019-02-22

dkrumm@reed.edu *Reed College, Portland, OR, United States*

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. 4 2019